Linker Dynamics in Allosteric Regulation of 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Science in Biochemistry

In the Department of Chemistry

At the University of Canterbury

By Jordyn Alissa Moore

March 2016

Abstract

The enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) catalyses the first step in the shikimate pathway, the condensation reaction between phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P); leading to production of the aromatic amino acids tryptophan (Trp), phenylalanine (Phe) and tyrosine (Tyr). DAH7PS enzymes are classified into two homologous families, type I and type II, differentiated by molecular weight and amino acid sequence. The type I DAH7PS enzymes are further sub-divided into type Iα and type Iβ, based on amino acid sequence. DAH7PS enzymes share a common TIM barrel fold, which comprises the key catalytic residues required for DAH7PS activity. The presence of recruited structural elements to the core TIM barrel fold, confers allosteric regulation to the DAH7PS enzymes.

The aim of this thesis is to explore the role of linker dynamics in allosteric regulation of the type Iβ DAH7PS enzyme, from the bacterium *Thermotoga maritima*. The hypothesis that the inter-domain linker determines molecular motion of the regulatory domains and consequently is responsible for allosteric regulation of the enzyme will be investigated.

Chapter two describes the functional and structural characterisation of *T. maritima* DAH7PS (*Tma*DAH7PS) and three *Tma*DAH7PS linker variants, each involving a single amino acid substitution within the linker sequence. The response of the *Tma*DAH7PS linker variants to inhibitors will be investigated and compared to the wild-type *Tma*DAH7PS enzyme (^{WT}*Tma*DAH7PS).

Chapter three considers the role of the inter-domain linker in molecular motion of the enzyme, linking protein dynamics to the allosteric regulation mechanism employed by *Tma*DAH7PS. The studies presented in this chapter focus on the ^{V65P}*Tma*DAH7PS linker variant, where a valine (Val) at position 65 has been substituted for proline (Pro). Comparisons to ^{WT}*Tma*DAH7PS are drawn.

Chapter four combines the results of chapters two and three to present a more comprehensive discussion of the role of linker dynamics in the allosteric regulation mechanism employed by *Tma*DAH7PS.

Chapter five provides a summary of the work presented in this thesis, considering future work that would further understanding of the role of linker dynamics in *Tma*DAH7PS allostery.

Acknowledgements

The culmination of this thesis represents a rich and rewarding experience that would not have been possible without the knowledge and support of a number of people. The first of whom is my supervisor Emily Parker, whose enthusiasm and dedication to science is inspirational. I am truly grateful for your continual guidance and encouragement. Thank you for motivating me to become a better scientist.

Thanks must also be extended to the Parker research group. The team culture is apparent, with members willing to share their time, advice and friendship. Thank you to Eric Lang, for your patient and informative introduction to computational biology, Nicky Blackmore, for guiding me through numerous experimental techniques, Leyla Bustamante, for your assistance with molecular biology, Fiona Given and Yu Bai, for your helpfulness with general laboratory queries, and Emma Livingstone, for critiquing various drafts over the duration of this research.

Finally, thank you to my husband and family, your support is invaluable.

Contents

Abbreviations

ACT domain regulatory domain

Ala alanine

***Ape*DAH7PS** *Aeropyrum pernix* DAH7PS

Arg arginine

ATP adenosine triphosphate

BTP 1,3-(tris(hydroxyl-methyl)-methylamino)propane

DAH7P 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate

DAH7PS 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase

D_{max} maximum particle dimension

DNA deoxyribonucleic acid

dNTP deoxyribonucleotide triphosphate

DSF differential scanning fluorimetry

E4P eyrthrose-4-phosphate

***Eco*DAH7PS** *Escherichia coli* DAH7PS

***Eco*DAH7PS-Phe** *Eco*DAH7PS regulated by Phe

EDTA ethylenediaminetetraacetic acid

Enz enzyme

EPSPS 5-enolpyruvyl shikimate 3-phospahte synthase

FPLC fast protein liquid chromatography

HIC hydrophobic interaction chromatography

HSP heat shock protein

IPTG isopropyl-1-thiol-β-D-galactopyranoside

ITC isothermal titration calorimetry

k_{cat} turnover number

K_d dissociation constant

K_m Michaelis constant

LB lysogeny broth

Leu leucine

***Lmo*DAH7PS** *Listeria monocytogenes* DAH7PS

Lys lysine

MD molecular dynamics

***Mtu*DAH7PS** *Mycobacterium tuberculosis* DAH7PS

NBD nucleotide binding domain

***Nme*DAH7PS** *Neisseria meningitidis* DAH7PS

NMR spectroscopy nuclear magnetic resonance spectroscopy

^{P68A}*Tma*DAH7PS *Tma*DAH7PS linker variant

PBC periodic boundary conditions

PCR polymerase chain reaction

PDB Protein Data Bank

PEP phosphoenolpyruvate

***Pfu*DAH7PS** *Pyrococcus furiosus* DAH7PS

Phe phenylalanine

Pi inorganic phosphate

PKA protein kinase A

PME particle mesh Ewald

***P*(r)** pair-distance distribution function

Pro proline

^{R64P}***Tma*DAH7PS** *Tma*DAH7PS linker variant

R_g radius of gyration

RMSD root mean square deviation

SAXS small angle X-ray scattering

SBD substrate binding domain

***Sce*DAH7PS** *Saccharomyces cerevisiae* DAH7PS

SDS-PAGE Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SEC size-exclusion chromatography

Ser serine

SOC super optimal broth with catabolite repression

TAE tris base-acetic acid-EDTA

TIM triosephosphate isomerase

***Tma*DAH7PS** *Thermotoga maritima* DAH7PS

TMD targeted molecular dynamics

Trp tryptophan

Tyr tyrosine

^{V65P}*Tma*DAH7PS *Tma*DAH7PS linker variant

Val valine

^{WT}*Tma*DAH7PS wild-type *Tma*DAH7PS

List of Figures

List of Tables

Chapter One

Introduction

1.1 The Shikimate Pathway

Proteins are essential biomolecules that control the chemistry of life. Organisms are able to generate the amino acids necessary for protein synthesis, via the reaction chemistry available to them. Animals are less adept at amino acid biosynthesis, in comparison to other eukaryotes or micro-organisms.[1] The aromatic amino acids tryptophan (Trp), phenylalanine (Phe) and tyrosine (Tyr) are among those amino acids unable to be biosynthesised by animals. In micro-organisms, plants, a few eukaryotes and discovered more recently, apicomplexan parasites, these important amino acids are produced by the shikimate pathway.[1-2]

The shikimate pathway acts to convert phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) to chorismate, via seven enzyme catalysed reactions (Figure 1.1).[3-4] Chorismate represents an important branch point in the pathway and is the precursor to the aromatic amino acids Trp, Phe and Tyr, as well as other aromatic metabolites; including folates, salicyclates and quinones.[5-6] As chorismate is a point of divergence for the pathway, this step is carefully regulated, allowing the formation of products to be precisely controlled.[7-8] Trp biosynthesis is achieved through conversion of chorismate to anthranilate in one branch of the pathway, while a second branch results in formation of both Phe and Tyr, via the precursor prephenate.[6-9] The location of prephenate in the reaction pathway represents another key, regulatory checkpoint.[10]

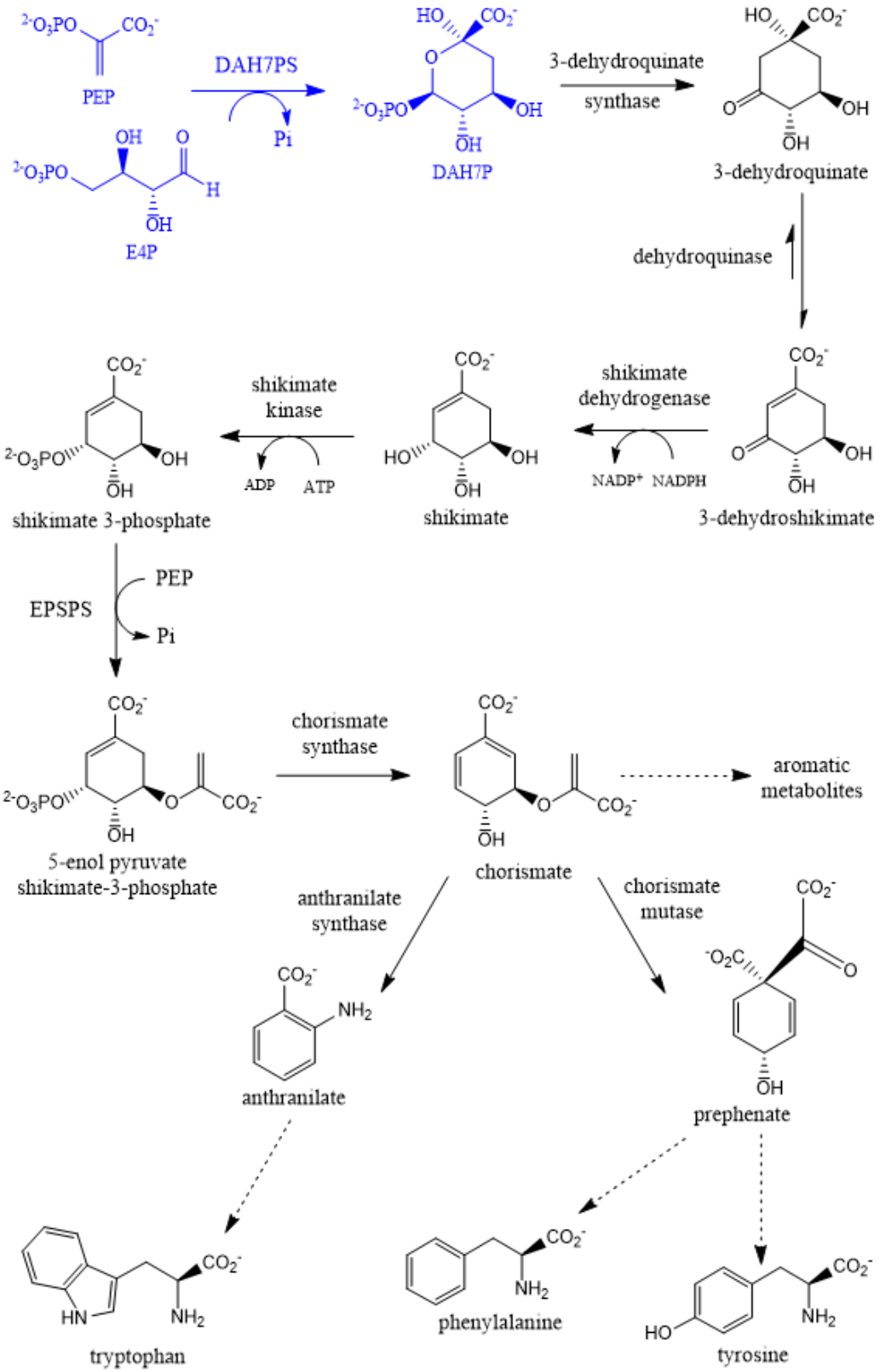

Figure 1.1: The seven enzyme catalysed reactions of the shikimate pathway, which lead to formation of chorismate, the precursor to the aromatic amino acids Trp, Phe and Tyr. Coloured in blue is the reaction catalysed by DAH7PS.

Animals have no equivalent reaction pathway and therefore the aromatic amino acids must be obtained through diet. The absence of the shikimate pathway from animal metabolism as well as multiple regulatory steps, present the shikimate enzymes as good targets for antimicrobial and herbicidal agents.[11] 5-enolpyruvyl shikimate 3-phospahte synthase (EPSPS), the sixth enzyme of the shikimate pathway, has already been shown to be inhibited by *N*-(phosphonomethyl)glycine, the active ingredient of the broad-spectrum herbicide glyphosate.[12-13] The success of glyphosate has prompted further investigation of the shikimate enzymes and their regulatory mechanisms, in the hope of uncovering more potent inhibitors for the pathway.[12-13]

1.2 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase

The first step of the shikimate pathway is catalysed by the enzyme 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS, EC 2.5.1.54).[14-15] DAH7PS enzymes catalyse the reaction between PEP and E4P to form 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH7P), with the release of inorganic phosphate (Pi), via an aldol like condensation mechanism (Figure 1.1). As the first enzyme of the shikimate pathway, DAH7PS represents a major control point for metabolism of the pathways aromatic products. The DAH7PS enzymes are often feedback regulated by the end products of the shikimate pathway, controlling DAH7PS activity in response to cellular requirements.[14,16-17] Multiple studies on DAH7PS enzymes from an array of organisms, has revealed numerous mechanisms exist to regulate DAH7PS activity.[18-19] DAH7PS enzymes have evolved a number of different allosteric control mechanisms to establish feedback regulation and exhibit varied responses to the aromatic amino acids Trp, Phe and Tyr, as a result.[18,20]

1.2.1 Catalytic Mechanism

The reaction mechanism for the conversion of PEP and E4P to DAH7P by the enzyme DAH7PS was initially predicted to undergo a ping-pong mechanism.[21] Later kinetic studies, with increased purity of enzyme, confirmed that the kinetic mechanism is actually ordered sequential; where PEP binds to the enzyme first, followed by E4P binding, loss of the phosphate group and finally DAH7P release.[22]

DAH7PS catalysis is predicted to occur in three separate steps (Figure 1.2). The first step is thought to be initiated by nucleophilic attack of C_3 of PEP, on C_1 of E4P; which is activated for attack by coordination to a divalent metal ion.[23-24] This step is stereospecific, and leads to formation of an oxocarbenium ion, which quickly reacts with water at the active site, to produce a bisphosphate hemiketal intermediate.[25] The tetrahedral intermediate is subjected to C-O bond cleavage, releasing phosphate and forming acyclic DAH7P.[23] Cyclisation of DAH7P is expected to occur after enzyme release, with greater than 99% adoption of the more stable pyranose form.[26]

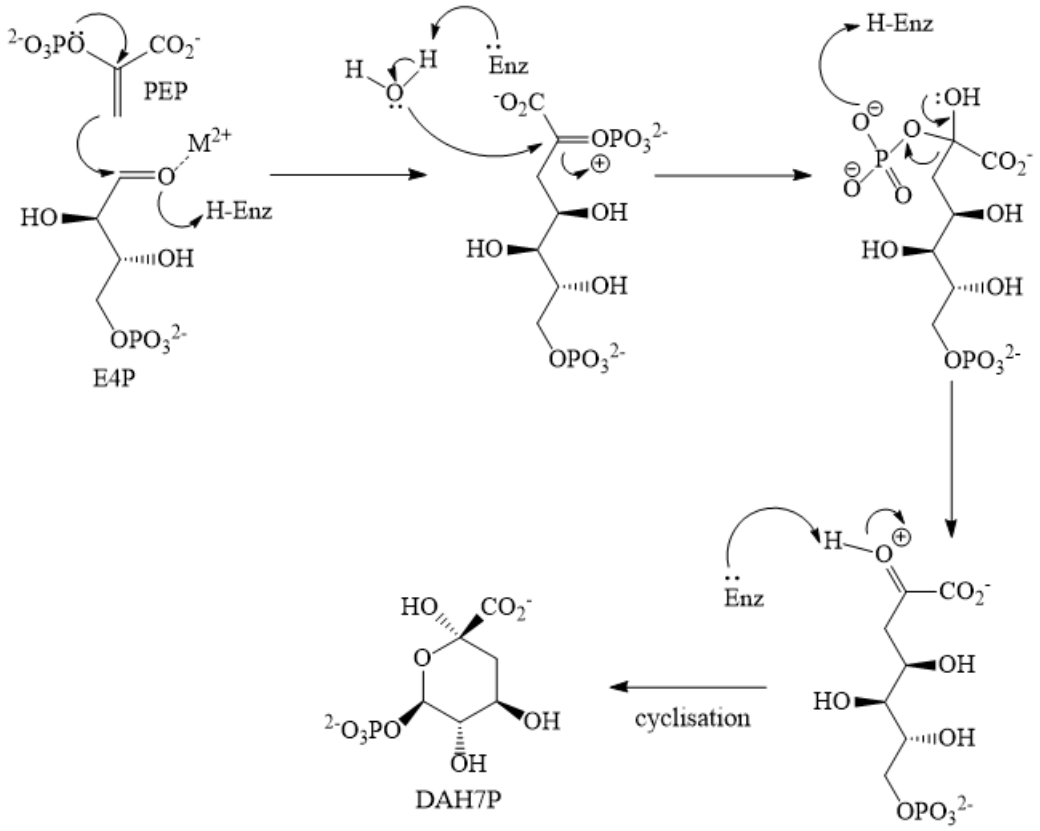

Figure 1.2: The proposed mechanism of PEP and E4P condensation to form DAH7P, catalysed by the enzyme DAH7PS.

Mechanistic detail for DAH7PS catalysis has been elucidated through several key labelling studies. Utilisation of deuterated[27] and fluorinated[28] PEP enabled the detection of stereospecific product formation by nuclear magnetic resonance (NMR) spectroscopy, confirming attack of the *si* face alkene functionality of PEP, to the *re* face aldehyde functionality of E4P.[25,29] In addition ^{18}O labelling exposed that the reaction proceeds via cleavage of the C-O bond of PEP, rather than the high energy O-P bond.[30] In order for this to occur, water would need to be involved in nucleophilic attack at position C_2 of PEP.[31] Indeed incorporation of the ^{18}O label in the DAH7P product was observed, corroborating the role of water in the catalytic mechanism of DAH7PS.

1.2.2 Classification of DAH7PS Enzymes

DAH7PS enzymes from different organisms have been classified into two distinct homologous families; designated type I and type II, based on molecular weight and amino acid sequence.[32-33] The type I enzymes are smaller in size, typically only 30-40 kDa, while the type II enzymes often have a molecular weight greater than 50 kDa. The sequence similarity between the two families is less than 10%.[34] The type I enzymes are further subdivided into type Iα and type Iβ, differentiated by amino acid sequence. There is approximately only 20% sequence identity between each subtype.[35] The type I DAH7PS enzymes are nearly always expressed in microbial organisms. However, the type II DAH7PS enzymes are expressed in both higher microbes and plants.[36-37]

The elucidation of a number of DAH7PS enzyme structures, suggests that the DAH7PS enzymes have actually been classified based on structural additions to the core TIM barrel fold.[38] The $(\beta\alpha)_8$ TIM barrel is the most commonly adopted protein fold and was first identified in triosephosphate isomerase.[39-40] The protein fold is common to many enzymes involved in aromatic amino acid metabolism and has been identified in the tertiary structure of all DAH7PS enzymes.[39-40]

It is hypothesised that many of the DAH7PS enzymes have arisen from gene duplication and gene fusion processes, due to the similarities between the DAH7PS types.[39] The conserved monomeric TIM barrel fold, binding sites for substrate PEP, and positioning of key active site residues, indicate a common ancestry between the DAH7PS types.[41] Differences between the DAH7PS enzymes involve extensions from the catalytic $(\beta\alpha)_8$ TIM barrel and oligomeric protein states, features both linked to regulatory function.[41] It is likely that divergent evolution from a common ancestor has led to the distinct DAH7PS types observed today.[41]

1.2.3 Structure of DAH7PS Enzymes

The structure of DAH7PS enzymes from each classification is presented, beginning with the smaller, type Iβ DAH7PS enzymes and progressing to the larger, type II DAH7PS enzymes.

Type Iβ DAH7PS Enzymes

The thermophilic archaea *Pyrococcus furiosus* expresses a single, unregulated type Iβ DAH7PS enzyme (*Pfu*DAH7PS) (Figure 1.3, left).[42] DAH7PS enzymes adopt varied quaternary structures dependent on the organism in which they are expressed. *Pfu*DAH7PS crystallises as a homotetramer, comprising two distinct interfaces (Figure 1.3, right).[43]

The structure of *Pfu*DAH7PS reveals the basic architecture, which all DAH7PS enzymes build upon, the $(\beta\alpha)_8$ TIM barrel.[43-44] The TIM barrel fold consists of eight α-helices and eight parallel β-strands. As is typical for enzymes with this structure, the active site for DAH7PS enzymes is located at the C-terminal end of the $(\beta\alpha)_8$ TIM barrel.[40] Active site geometry is facilitated by the extended β-α loops at the C-terminal, which provide the active site residues required for DAH7PS catalysis.[43] The absence of an extension to the core TIM barrel of *Pfu*DAH7PS, results in the enzyme being unregulated. No inhibition in the presence of any of the downstream products of the shikimate pathway has been observed for any DAH7PS enzyme with this architecture.[44-45]

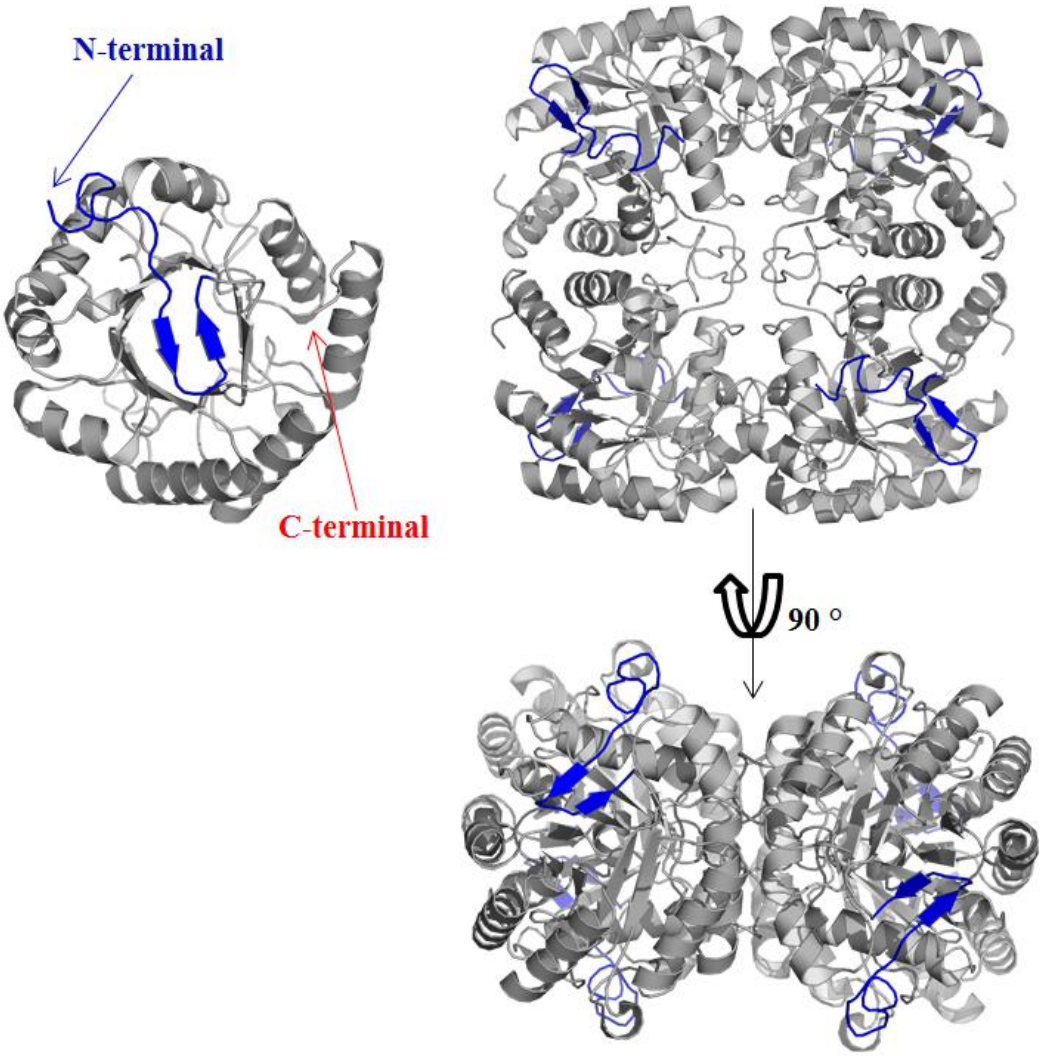

Figure 1.3: Monomeric (left) and tetrameric (right) structures of *Pfu*DAH7PS (PDB code 1ZCO), with the $(\beta\alpha)_8$ TIM barrel fold coloured in grey and the N-terminal extension coloured in blue. The tetrameric structure has been rotated about its diagonal axis, highlighting the protein interfaces.

The thermophilic bacterium *Thermotoga maritima* also expresses a single, type Iβ DAH7PS enzyme (*Tma*DAH7PS). However, unlike *Pfu*DAH7PS, the $(\beta\alpha)_8$ TIM barrel of *Tma*DAH7PS is adorned with a N-terminal regulatory domain. Each of the four monomeric units of *Tma*DAH7PS comprises a core TIM barrel fold; connected by a flexible linker region to a N-terminal regulatory domain, which adopts an ACT-like fold (Figure 1.4, left).[37]

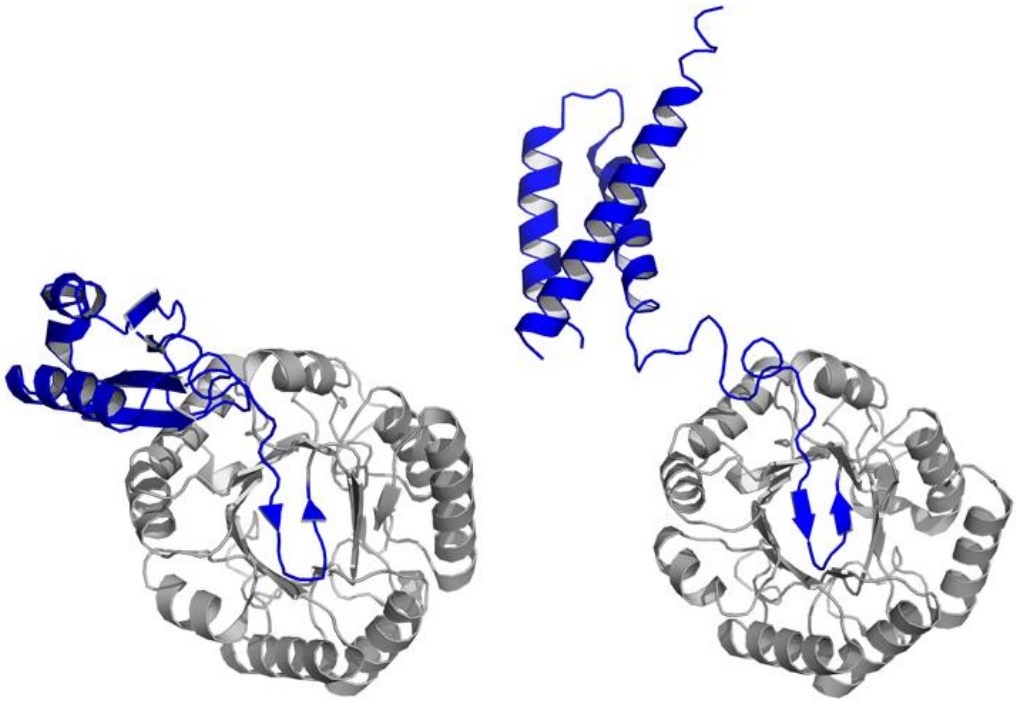

Figure 1.4: Monomeric structures of *Tma*DAH7PS (PDB code 1RZM) (left) and *Lmo*DAH7PS (PDB code 3NVT) (right), with the $(\beta\alpha)_8$ TIM barrel fold coloured in grey and the N-terminal extension coloured in blue.

The ACT domain was named for the first three proteins it was discovered in: aspartate kinase, chorismate mutase and prephenate dehydrogenase, or tyrosine A; and functions as a ligand binding site for regulatory control of metabolic pathways, where the ligand bound is often the end product of the reaction series.[46-47] Pervasiveness of the ACT domain in nature, suggests an early evolutionary recruitment for allosteric regulation of metabolic pathways.[46-47]

The crystal structure of *Tma*DAH7PS is also homotetrameric. The four regulatory domains of *Tma*DAH7PS form the allosteric sites for the enzyme and function in a unique mechanism of allosteric regulation.[48] *Tma*DAH7PS activity is regulated predominantly by the amino acid Tyr. In the absence of Tyr, the regulatory domains orient away from one another (open conformation), enabling substrates to access the active site however, when Tyr is bound a dramatic reorganisation of the regulatory domains occurs (closed conformation) (Figure 1.5).[48]

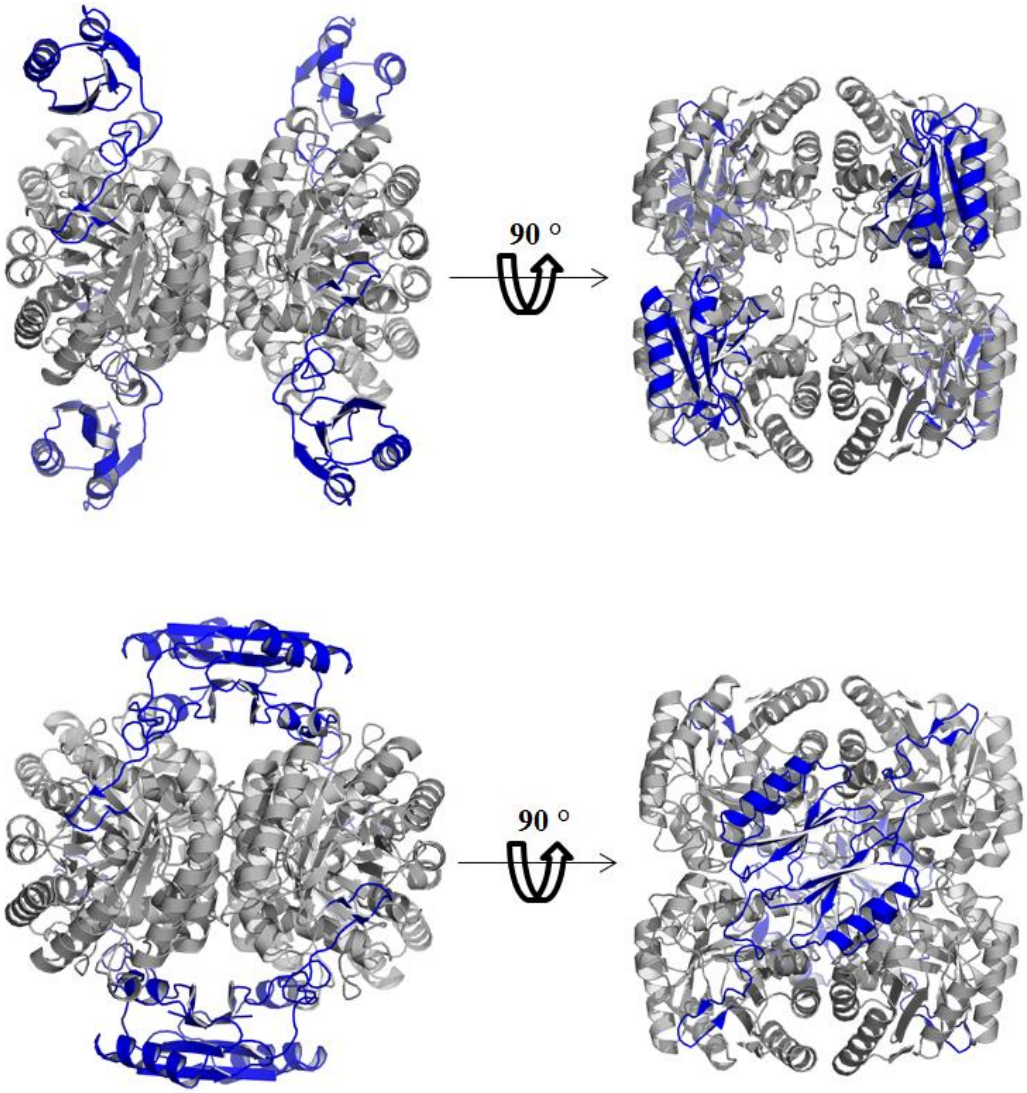

Figure 1.5: Tetrameric structures of *Tma*DAH7PS in the open (above) and closed conformation (PDB code 3PG9) (below) oriented to enable the N-terminal regulatory domains to be viewed. For each structure the $(\beta\alpha)_8$ TIM barrel fold is coloured grey and the N-terminal extension coloured in blue.

In the presence of Tyr the N-terminal regulatory domains undergo a large rotation, with diagonally opposite regulatory domains moving toward the central axis of the tetramer.[48] In their altered position the regulatory domains form new contacts with one another and the catalytic $(\beta\alpha)_8$ TIM barrel core.[48] The inhibitor bound conformation places the regulatory domains directly above the TIM barrel fold, decreasing the ability of any substrate to reach the active site and greatly reducing catalytic activity of the enzyme.

DAH7PS from the bacterium *Listeria monocytogenes* (*Lmo*DAH7PS) represents another type Iβ DAH7PS enzyme but is regulated by chorismate and prephenate, the substrate and product of the reaction catalysed by chorismate mutase.[49-50] The catalytic domain of *Lmo*DAH7PS adopts the common TIM barrel fold. However, an N-terminal extension with a high degree of sequence homology to the chorismate mutase enzyme is present (Figure 1.4, right). The chorismate mutase-like regulatory domains are positioned above and below the tetrameric assembly for *Lmo*DAH7PS.[50]

DAH7PS enzymes of this form can also be regulated by chorismate and prephenate via a C-terminal, instead of an N-terminal extension. The main difference between the protein architectures is the location of the regulatory domains and domain linkers.[9,50] The inter-domain linker, which connects the regulatory domain to the core TIM barrel, enables the chorismate mutase-like domain to be positioned at either terminus of the TIM barrel and confer allosteric regulation to the enzyme.

For type Iβ DAH7PS enzymes it is the interaction between diagonally opposite and not adjacent monomers of the tetramer that confers allosteric regulation.[34] This suggests that homomeric quaternary structure is essential for the allosteric regulation mechanisms employed by the type Iβ DAH7PS enzymes.[34]

Type Iα DAH7PS Enzymes

In contrast to the organisms that express type Iβ DAH7PS enzymes, organisms which encode type Iα DAH7PS enzymes often express multiple isozymes that are feedback regulated by Phe, Tyr or Trp. The *Escherichia coli* genome encodes three type Iα DAH7PS isozymes (*Eco*DAH7PS). Each isozyme exhibits specific sensitivity to a single aromatic amino acid.[51-52] The Phe responsive *Eco*DAH7PS (*Eco*DAH7PS-Phe) is among the best characterised DAH7PS enzymes and was the first DAH7PS structure to be solved by X-ray crystallography.[53]

The overall structure of *Eco*DAH7PS-Phe is similar to the unregulated type Iβ DAH7PS enzymes. However, *Eco*DAH7PS-Phe has two modifications to the core $(\beta\alpha)_8$ TIM barrel. The first is a helical N-terminal extension and the second is a β-hairpin fold, present at the catalytic core (Figure 1.6, left).[53] The β-hairpin fold interacts directly with the N-terminal extension to form a pocket, where Phe can bind.[9]

Figure 1.6: Monomeric structure of *Eco*DAH7PS-Phe (PDB code 1QR7) (left), *Sce*DAH7PS (PDB code 1OFP) (centre) and *Nme*DAH7PS (PDB code 4HSN) (right), with the $(\beta\alpha)_8$ TIM barrel fold coloured in grey, the N-terminal extension coloured in pink and the β-hairpin fold coloured in yellow.

The yeast *Saccharomyces cerevisiae* expresses two type Iα DAH7PS isozymes (*Sce*DAH7PS), one inhibited by Phe and the other by Tyr.[24,39] The presence of only two isozymes for *S. cerevisiae* is likely due to the fact that both isozymes can be inhibited by Trp, in addition to either Phe or Tyr.[54] Both *Sce*DAH7PS isozymes have extended $(\beta\alpha)_8$ TIM barrels, with a similar architecture to *Eco*DAH7PS-Phe (Figure 1.6, centre).[24] For both *Eco*DAH7PS-Phe and *Sce*DAH7PS the N-terminal domain from one monomer interacts with the N-terminal domain of another monomer, to form two tight dimers that compose the tetramer (Figure 1.7).[24]

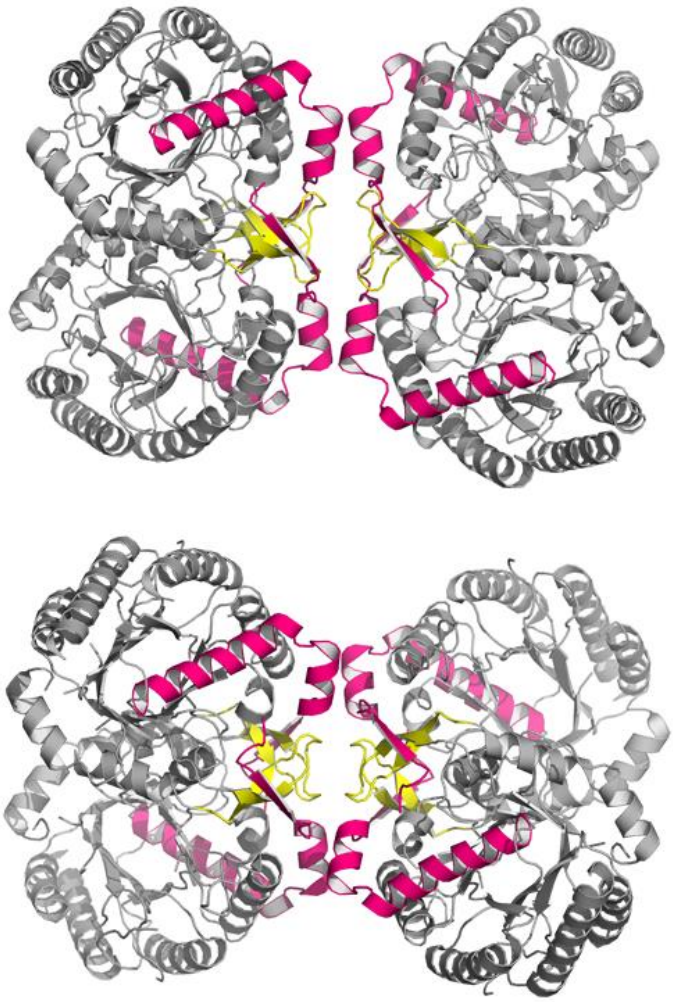

Figure 1.7: Tetrameric structure of *Eco*DAH7PS-Phe (above) and *Sce*DAH7PS (below). For each monomeric unit the $(\beta\alpha)_8$ TIM barrel fold is coloured in grey, the N-terminal extension in pink and the β-hairpin fold in yellow.

The *Neisseria meningitidis* genome encodes a single DAH7PS enzyme (*Nme*DAH7PS), which is inhibited by all three aromatic amino acids but shows the greatest level of inhibition in response to Phe.[38,55] This is in direct contrast to the other type Iα DAH7PS enzymes, which are primarily sensitive to only a single amino acid. The structure of *Nme*DAH7PS also shows many similarities to *Eco*DAH7PS-Phe; crystallising as a homotetramer, with monomer subunits comprising a $(\beta\alpha)_8$ TIM barrel core with an N-terminal extension (Figure 1.6, right).

Type II DAH7PS Enzymes

The best studied bacterial type II DAH7PS enzyme is the DAH7PS enzyme expressed by *Mycobacterium tuberculosis* (*Mtu*DAH7PS). *Mtu*DAH7PS is composed of four $(\beta\alpha)_8$ TIM barrels with N-terminal extensions that associate to form a homotetramer.[41] The catalytic core of *Mtu*DAH7PS is relatively similar in structure to that of the type I DAH7PS enzymes and has a comparable active site. However, the β2-α2 loop, which functions in substrate recognition, is extended (Figure 1.8).[41]

Figure 1.8: Monomeric structure of *Mtu*DAH7PS (PDB code 2B7O), with the TIM barrel fold coloured in grey, the N-terminal extension coloured in pink and the extended β2-α2 loop coloured in cyan.

Two distinct Phe and one Trp binding site, which are related synergistically, have been identified for *Mtu*DAH7PS.[56] Trp binding occurs at site one, positioned at the tetramer interface, while Phe binds at site two, located at the dimer interface.[11,56] Site three is close to the dimer interface and is also found to bind Phe. However, Tyr is bound preferentially. In the presence of Phe, Tyr binding only occurs at site three.[57]

The sites required to elicit the inhibitory response for *Mtu*DAH7PS are sites one and two, verified by mutagenesis studies on site three, which concluded site three was not necessary for inhibition by Trp and Phe but enabled additional response to Tyr.[11,56] The Phe and Trp binding sites communicate occupancy with one another and the active site.[56] The allosteric response is precisely controlled by the number of sites that are occupied and the amino acids that occupy each site.[11] A comparative analysis of the *Mtu*DAH7PS ligand bound and ligand free structures indicates there is no average conformational change associated with ligand binding.[56-57]

1.3 Enzyme Regulation

Metabolism encompasses all chemical transformations that occur within the cells of living organisms.[58] Metabolic pathways are formed by a series of interconnected biochemical reactions that are governed by enzymes. A metabolic pathway converts an initial metabolite to a specific biomolecule via a sequence of chemical reactions, each catalysed by an enzyme; where the product of the former enzyme becomes the substrate for the next, within the reaction series.[58-59] The presence of regulatory enzymes at key junctions of the metabolic pathway, which are able to sense the metabolic requirements of the cell and adjust their catalytic activity in response, controls biomolecule formation.[59-60]

Metabolic flux can be controlled at the level of the enzyme, through modulation of enzyme activity, or at the level of the gene, where transcriptional activity of the gene encoding the enzyme is regulated. This presents several mechanisms to control enzyme catalysis, including transcription of DNA into mRNA,[61] translation of mRNA into amino acids, post-translational protein modifications,[62] compartmentalisation and localisation of the enzyme,[63] and feedback inhibition mechanisms, such as allostery.[64]

1.3.1 Allosteric Regulation

Allosteric regulation is crucial in the management of metabolic processes, signal transduction events and gene regulation.[65] Allostery is defined by specific ligand binding that influences the function of a remote site of a protein. The distal ligand binding site is the allosteric site and the ligand that binds to the allosteric site is termed the allosteric effector.[66-67] The allosteric response can be initiated through ligand binding, as is the case for DAH7PS enzymes but also through covalent modification or mutation at the allosteric site.[67-68] The effect of binding an allosteric effector to the allosteric site, is transmitted through a protein, resulting in a change in protein function. Allosteric effectors can either increase or decrease protein function and are referred to as either allosteric activators or inhibitors respectively.[66-67] Changes in protein function can include alterations in catalytic activity and/or ligand affinity (Figure 1.9).[66-67]

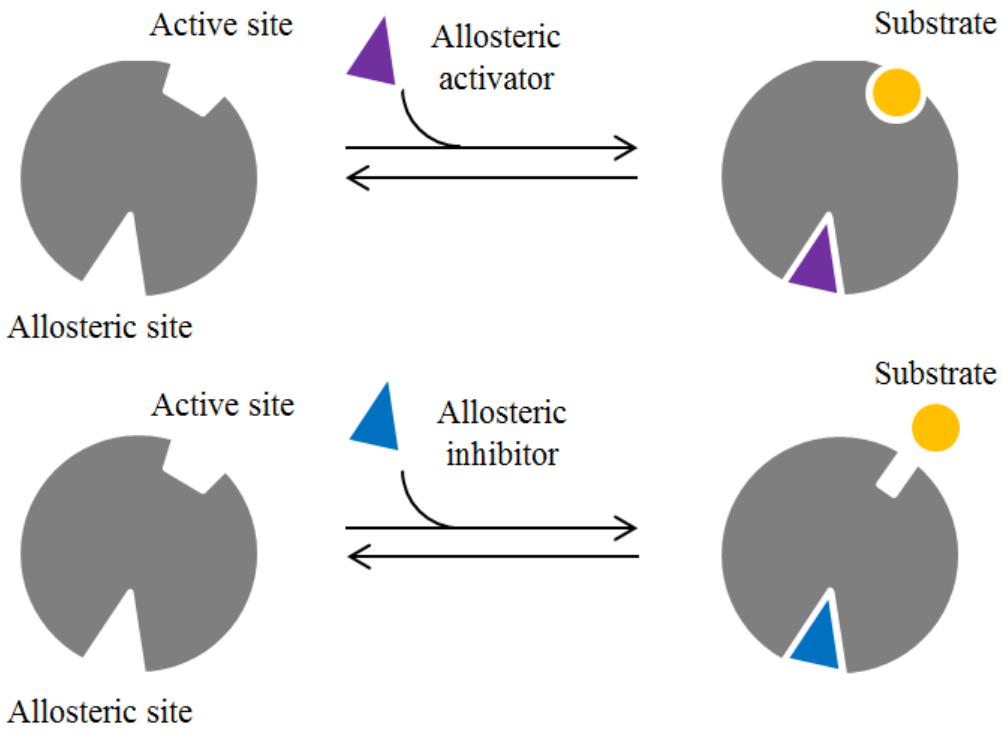

Figure 1.9: Schematic representation of an allosterically regulated enzyme, shown in grey, where substrate affinity is altered upon allosteric effector binding. Binding of an allosteric activator to the allosteric site increases substrate affinity (above), while binding of an allosteric inhibitor to the allosteric site decreases substrate affinity (below).

History of Allosteric Regulation

The concept of allostery was first established by Changeux, who identified that two distinct binding sites within a protein, assocciated with different ligands, could interact without proximity to one another.[66] This led to the proposal of two models of allostery. The Monad-Wyman-Changeux (MWC) model,[69] where ligand binding shifts the equilibrium between two distinct conformational states, and the Koshland-Nemethy-Filmer (KNF) model,[70] where proteins are considered to be flexible species, which undergo an induced-fit-like alteration in protein structure, in response to ligand binding.[66,71] Both models can be used to describe allosteric regulation.

In the past, allosteric regulation of protein function was considered to result from changes in protein structure, induced by ligand binding.[65,67] Proteins that underwent large conformational changes in response to ligand binding were consequently used as model systems to describe allostery. Haemoglobin was one of the first proteins discovered to be regulated by an allosteric mechanism.[66-71] Composed of four subunits, haemoglobin undergoes detectable changes in conformation upon ligand binding.[67] The clearly defined "T" (for tense) and "R" (for relaxed) states are easily identified and initial understanding of allostery came from observed differences in haemoglobin structure, resulting from allosteric effector binding.[66, 70] As understanding of allostery progressed, it was realised that allostery is not limited to multimeric proteins that undergo changes in conformation but can also occur in monomeric proteins.[67]

Following this, it was later realised that allosteric regulation can take place even in the absence of gross conformational change. It is now accepted that conformational change in response to ligand binding is only one mechanism of allosteric regulation and that allostery is also associated with changes in protein dynamics and conformational flexibility.[66] Thus the molecular communication networks between the allosteric and active sites can be mediated by either

structural changes or alterations in relative protein flexibility, exemplifying the role of both enthalpic and entropic factors in allostery.[67,72]

1.3.2 Regulation of DAH7PS Enzymes

DAH7PS enzymes display a remarkable diversity of allosteric regulation mechanisms, which are necessary for precise control of metabolic flux. The evolution of allosteric control for DAH7PS enzymes is postulated to have occurred via domain recruitment, with the addition of a regulatory domain to the $(\beta\alpha)_8$ TIM barrel core.[16,48] Initial studies on the unregulated DAH7PS enzymes helped to confirm that structural extensions to the $(\beta\alpha)_8$ TIM barrel are responsible for allosteric control of DAH7PS enzymes. It is possible that the unregulated DAH7PS enzymes are more primitive than their regulated counterparts and regulation is achieved at a different step of the pathway, or even at the level of transcription.[44]

Type Iβ DAH7PS Enzymes

Allosteric regulation of *Tma*DAH7PS is perhaps one of the best understood type Iβ DAH7PS regulation mechanisms that involve a conformational change upon ligand binding. *Tma*DAH7PS activity is regulated by Tyr and to a much lesser extent by Phe.[9,7,48] In the absence of Tyr, the regulatory domains orient away from one another, enabling PEP and E4P to enter the active site. However, when Tyr is present the regulatory domains undergo a dramatic reorganisation; altering the affinity of the protein for its natural substrates and leading to inhibition of the enzyme.[48]

A dynamic equilibrium between the active and inactive forms of *Tma*DAH7PS appears to exist, where the presence of Tyr alters the position of the equilibrium. In the absence of Tyr, the open conformation of *Tma*DAH7PS is favoured and the enzyme is active but upon Tyr binding the equilibrium is shifted, and the enzyme adopts a closed conformation that is inactive. If this is correct, Tyr is selectively

binding to a pre-existing but less populated sub-state of the protein ensemble; leading to a shift in the observed conformer, as a mechanism of allosteric regulation.[48]

A large degree of structural similarity exists between the type Iβ DAH7PS enzymes *Lmo*DAH7PS and *Tma*DAH7PS. Analysis of the structural features of each enzyme, led to the prediction that the mechanism of allosteric regulation for *Lmo*DAH7PS, similar to *Tma*DAH7PS, might involve a conformational adjustment of the regulatory domains in response to inhibitor binding.[18,49] The conformational change has been described for *Tma*DAH7PS but not yet for *Lmo*DAH7PS. If the mechanism of allosteric regulation for the chorismate/prephenate regulated DAH7PS enzymes involves a comparable change in conformation to *Tma*DAH7PS, the similarity in structure and movement of the regulatory domains would indicate a conserved mechanism of allosteric regulation among the type Iβ DAH7PS enzymes; despite overall differences in the regulatory domains and inhibitor species.[9]

*Lmo*DAH7PS is allosterically regulated by chorismate and prephenate, upon ligand binding to the chorismate mutase-like regulatory domain.[9,49] It is possible that the chorismate mutase-like regulatory domain, has arisen evolutionarily through gene duplication and gene fusion, to confer allosteric regulation.[49, 73] Gene duplication and gene fusion events would be an efficient use of genetic resources and are likely, considering other DAH7PS enzymes appear to have evolutionarily recruited their regulatory domains. This hypothesis is based on the observation that the chorismate mutase-like domain exhibits strong feedback inhibition by prephenate.[73] It seems likely that an unregulated DAH7PS enzyme is the common ancestor of DAH7PS enzymes, and that allosteric regulation may be the evolutionary link between all DAH7PS enzymes.[73]

Type Iα DAH7PS Enzymes

The allosteric regulation mechanisms of the type Iα DAH7PS enzymes are more subtle than those of the type Iβ DAH7PS enzymes, which involve larger

conformational adjustments of discrete regulatory domains. In comparison, the type Iα DAH7PS enzyme regulation mechanisms involve much smaller conformational adjustments that lead to enzyme inhibition.

It has been hypothesised that the type Iα DAH7PS isozyme *Eco*DAH7PS-Phe is allosterically regulated via a series of conformational changes, which propagate through the enzyme to the active site, upon binding of Phe. It was reported that the series of conformational changes alter positioning of the β2-α2 loop, which is required for substrate interaction.[74] It is predicted that the orientation of PEP is altered in the inhibited conformation, preventing the necessary interactions required for catalytic activity. However, the altered positioning of PEP has only been observed in one study.[74] *Sce*DAH7PS has a very similar structure to *Eco*DAH7PS-Phe, with slight modifications which form an allosteric site for Tyr binding. As would be expected from their structures both enzymes employ a similar mechanism of allosteric regulation.[39]

*Nme*DAH7PS is inhibited by all three aromatic amino acids. However, the sensitivity of the enzyme to Phe, Tyr and Trp differs. *Nme*DAH7PS shows the greatest level of inhibition in response to Phe but the response to Trp is significantly less than the other aromatic amino acids.[38] There is no significant conformational change for *Nme*DAH7PS observed upon inhibitor binding, so the mechanism of allosteric regulation likely involves subtle changes in protein dynamics.[38]

Type II DAH7PS Enzymes

There is much interest in the regulatory mechanisms of the type II DAH7PS enzymes because of their potential as drug targets, in the pathogenic organisms in which they are expressed.[41] *M. tuberculosis* is the causative agent of tuberculosis and gene disruption studies have revealed that the shikimate pathway is essential for *M. tuberculosis* viability, prompting thorough examination of this DAH7PS enzyme.[41]

*Mtu*DAH7PS reveals a complex allosteric regulation mechanism unlike that identified for any other DAH7PS enzyme.[75] *Mtu*DAH7PS activity is inhibited by a combination of Phe, Tyr and Trp, in a complex synergistic regulation mechanism.[9,11,56] Interestingly the inhibitory effect of Phe, Tyr and Trp in combination, is much greater than that observed for any single inhibitor in isolation.

Synergy is observed in particular between Phe and Trp, which exhibit the greatest level of inhibition compared to any other combination.[41,57] Subsequent addition of Tyr to Phe and Trp, markedly increases the inhibitory response, with Tyr binding resulting in complete loss of catalytic activity.[11] It is predicted that a synergistic DAH7PS may have evolved, to allow more precise control over regulation processes, in response to metabolic requirements.[11] The synergistic response of *Mtu*DAH7PS enables highly specific production of the shikimate amino acids.[57]

The mechanism of allosteric regulation for *Mtu*DAH7PS, likely involves an alteration in the dynamic properties of the protein; in response to synergistic binding of Phe and Trp.[56-57] Analysis of temperature factors over the course of a molecular dynamics simulation suggests that binding of Phe and Trp, increases the flexibility of the β2-α2 loop. Increased flexibility of the β2-α2 loop is hypothesised to increase the entropy penalty of substrate binding, making it more difficult to bind substrate and leading to enzyme inhibition.[9,56] The proposed mechanism of inhibition supports the theory that the β2-α2 loop is conserved across DAH7PS types because it plays an important role in allosteric regulation of DAH7PS enzymes.[9]

The unusual approach to allosteric regulation demonstrated by *Mtu*DAH7PS provides some insight into the different mechanisms of allostery that can evolve. The genome of *M. tuberculosis* codes for only a single DAH7PS enzyme, which may be the reason for the need to evolve a highly sensitive enzyme that is able to be regulated effectively, by multiple aromatic amino acids.[11] Other enzymes known to display synergistic allostery, such as glutamine

phosphoribosylpyrophosphate amidotransferase from *Bacillus subtilis*, are also involved in the first step of the biosynthetic pathway in which they function.[11] Supporting the idea that synergistic allostery is employed when multiple end products are produced and regulation of metabolic control is particularly important.

Summary of DAH7PS Enzyme Regulation

DAH7PS enzymes display an immense variety of allosteric regulation mechanisms that are facilitated by the presence of discrete regulatory domains or smaller extensions to the core, catalytic $(\beta\alpha)_8$ TIM barrel. It has been proposed that the mechanism of allosteric regulation defines each of the DAH7PS enzyme types. This continues to be supported, as additional DAH7PS structures are reported and the enzymes regulatory mechanisms discovered.

Comparison of the DAH7PS enzymes exposes many structural similarities. However, differences in features related to allosteric regulation are apparent. Small alterations to the allosteric binding pocket characterise the type Iα DAH7PS enzymes, while the position of discrete regulatory domains define the type Iβ DAH7PS enzymes.[9] It seems likely that DAH7PS enzymes have evolved from a primitive, unregulated DAH7PS enzyme and that over the course of evolution, some members of the DAH7PS family have acquired extra domains for the purpose of feedback regulation. Regardless, the DAH7PS enzymes exhibit a wide range of allosteric regulation mechanisms, which incorporate both conformational and dynamic changes to facilitate specific production of the amino acids of the shikimate pathway.

1.4 Objectives of Thesis

The overall aim of this research is to explore allosteric regulation of *Tma*DAH7PS, by investigating the role of the inter-domain linker in the enzymes mechanism of allostery. To date there has been no investigation into the flexibility of the *Tma*DAH7PS linker region, nor has the dynamic role the linker plays in the conformational change upon Tyr binding been studied. Tyr binding potentially perturbs linker dynamics, facilitating allosteric propagation.[76] If this is the case, particular amino acids of the linker region may be conserved at set positions, to ensure transition through the propagation pathway upon Tyr binding.[76] Thus mutation of key amino acids of the linker region, may affect protein function, due to either altered flexibility or disruption of the allosteric propagation pathway; enabling the function of the linker sequence to be examined.[76]

The studies presented in this thesis investigate the following aims:

- Kinetically and structurally characterise *Tma*DAH7PS and several variant *Tma*DAH7PS enzymes, to examine the role of the linker sequence in the mechanism of allosteric regulation by Tyr.

- Explore protein dynamics of *Tma*DAH7PS and variant *Tma*DAH7PS enzymes, via molecular dynamics simulations, to determine if DAH7PS allosteric propagation pathways can be predicted *in silico*.

- Discuss the results from the biochemical characterisation of *Tma*DAH7PS and the *Tma*DAH7PS linker variants, in combination with the *in silico* results, to unravel the dynamic nature of the allosteric regulation mechanism employed by *Tma*DAH7PS, upon Tyr binding.

Chapter Two

Biochemical Characterisation of *Thermotoga maritima* DAH7PS and Three Linker Variants

2.1 Introduction

Hyperthermophiles are organisms that thrive in temperatures exceeding 60 °C and have been discovered in geothermally heated environments, such as deep-sea hydro-thermal vents.[42] Most hyperthermophiles belong to the domain archaea. However, bacteria classified in the order Thermotogales are also hyperthermophilic.[42,77] *T. maritima* is a member of the order Thermotogales, and was first discovered in the geothermally heated sediment off the coast of Vulcano, Italy. *T. maritima* is able to grow at temperatures between 55-90 °C, although the optimal temperature for growth is 80 °C.[78]

Hyperthermophiles are considered to be the earliest known ancestor of all extant life on Earth. Thermotogae have been the subject of numerous evolutionary studies, as the phylum is among the deepest branching bacterial species.[79-80] The *T. maritima* genome encodes a single type Iβ DAH7PS enzyme, and it is anticipated that investigation of this enzyme, will provide insight into the divergence of the DAH7PS enzyme classes and their distinct mechanisms of regulation.

The allosteric regulation mechanism of *Tma*DAH7PS involves gross conformational change, associated with Tyr binding. Conformational adjustment of the regulatory domains is facilitated by the flexible inter-domain linker, which comprises residues 64-80, and connects the catalytic TIM barrel to the N-terminal regulatory domain of each protein chain (Figure 2.1).

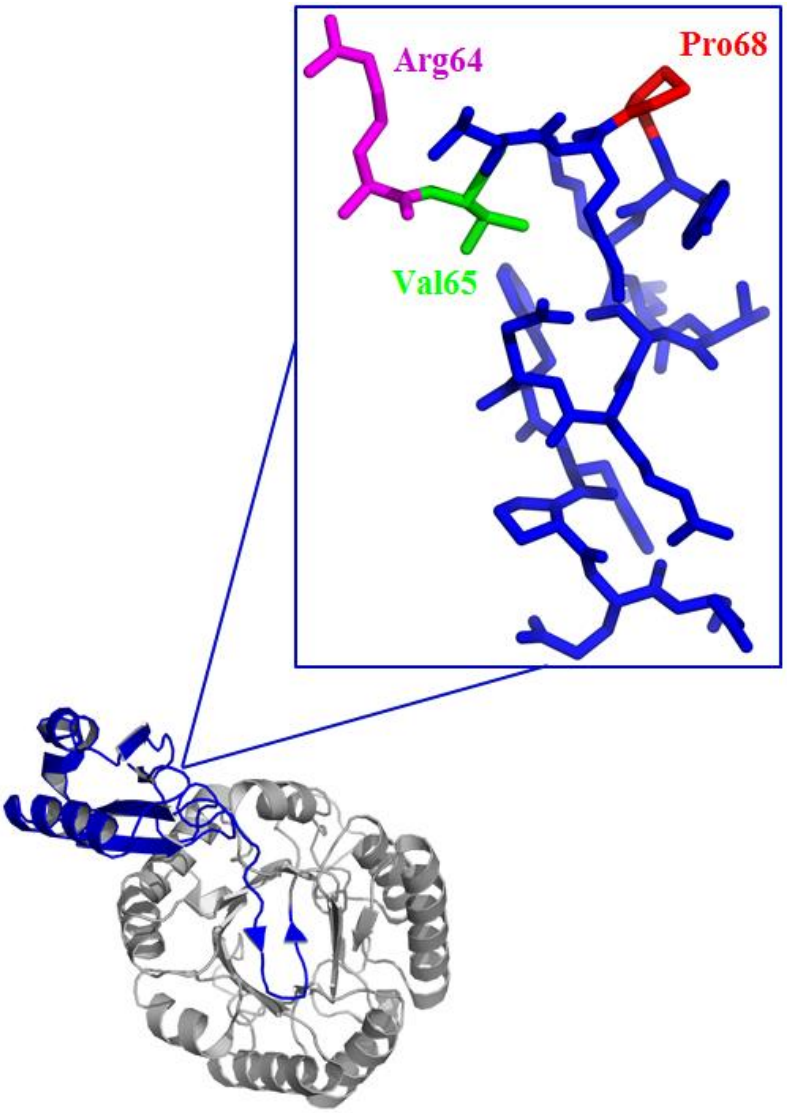

Figure 2.1: Monomeric structure of *Tma*DAH7PS, with the $(\beta\alpha)_8$ TIM barrel fold coloured in grey and the N-terminal extension coloured in blue. Highlighted is the inter-domain linker, with the residues arginine 64 (Arg64), valine 65 (Val65) and proline 68 (Pro68) labelled.

To transition from the open to closed conformation of *Tma*DAH7PS, the inter-domain linker must adopt altered conformations, in order to accommodate movement of the regulatory domains. Overlay of the open and closed monomeric structures of *Tma*DAH7PS, accentuates the altered conformation of the linker, for the open and closed enzyme forms (Figure 2.2).

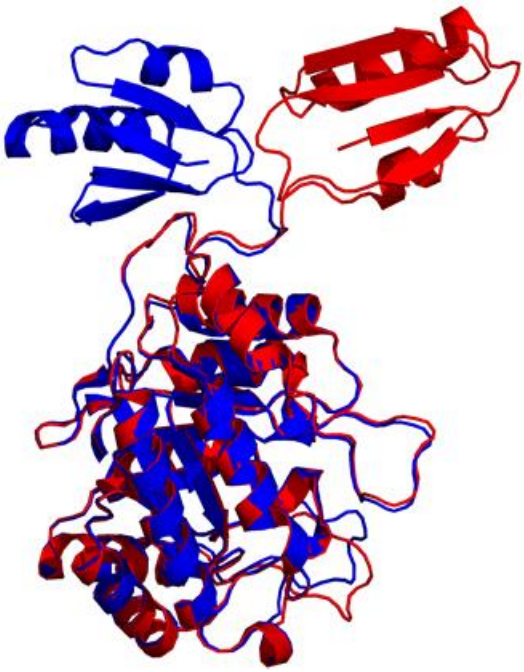

Figure 2.2: Monomer overlay of the open and closed conformations of *Tma*DAH7PS. The open, ligand free conformation, is coloured blue and the closed, Tyr bound conformation, is coloured red.

The movement of the regulatory domains in response to ligand binding necessitates some part of the linker region to be flexible. The allosteric regulation mechanism also requires the allosteric signal to be transmitted from the site of ligand binding, to the active site of the protein. It is possible that the amino acids of the linker sequence, encode an allosteric propagation pathway that enables the allosteric signal to be communicated to the active site for *Tma*DAH7PS.[81]

Characterisation of *Tma*DAH7PS has been reported previously[82] but is presented here, in combination with three new linker variants, to enable direct structural and functional comparisons to be made. This chapter describes the expression, purification and biochemical characterisation of wild-type *Tma*DAH7PS; herein described as ^{WT}*Tma*DAH7PS, along with three *Tma*DAH7PS linker variants.

2.2 Selection of *Tma*DAH7PS Linker Variants

With consideration to the inherent flexibility the linker must have, along with the possibility of a pre-encoded allosteric propagation pathway, the individual amino acids of the linker sequence were examined, to determine which amino acids are important for allosteric regulation of *Tma*DAH7PS. A sequence alignment was performed, to determine the amino acids of the linker, required for linker functionality (Figure 2.3). The residues valine 65 (Val65), tyrosine 69 (Tyr69), lysine 70 (Lys70) and serine 73 (Ser73) exhibited a high level of conservation among the type Iβ DAH7PS enzymes.

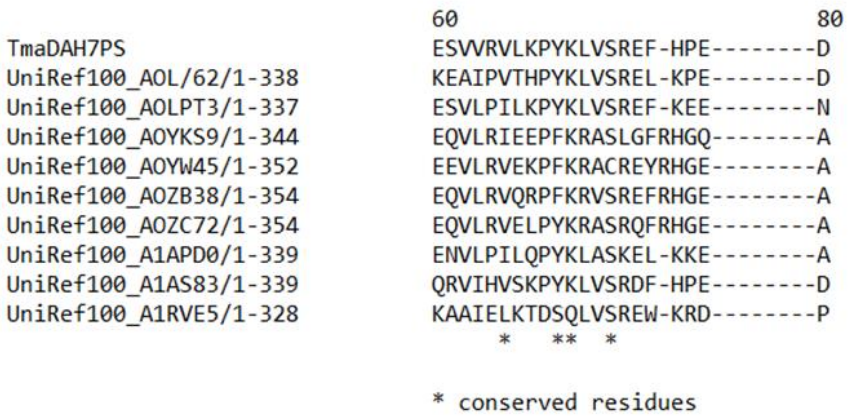

Figure 2.3: Clustal Ω sequence alignment of the type Iβ DAH7PS sequences. Of the full sequence alignment, only the residues of the linker, for the first ten type Iβ DAH7PS sequences are shown.

Linker variants were selected, to explore the effects of both decreased and increased linker flexibility. The first linker variant, substituted Val at position 65, for proline (Pro), generating an enzyme with reduced linker flexibility (^{V65P}*Tma*DAH7PS). ^{V65P}*Tma*DAH7PS was selected for two key reasons; first, because the high conservation score indicates that Val65 is conserved for functionality, amongst DAH7PS enzymes; second, a comparison of the ligand free and ligand bound crystal structures for *Tma*DAH7PS, demonstrated a large variation in the phi (φ) and psi (ψ) angles, of Val65. The difference in φ and ψ angles for Val65, was first reported by Cross[82] but has been investigated more

thoroughly here, and is presented in chapter three, within the targeted molecular dynamics section.

Following preliminary studies on ^{V65P}*Tma*DAH7PS, two additional linker variants were selected. The second linker variant substitutes Arg64 for Pro (^{R64P}*Tma*DAH7PS), and was selected for similar reasons to ^{V65P}*Tma*DAH7PS. Unlike position 65 however, position 64 is oriented away from the site of Tyr binding. The third linker variant substitutes Pro68 for alanine (Ala) (^{P68A}*Tma*DAH7PS), and was chosen to increase flexibility of the inter-domain linker. Pro68 is positioned at the centre of the most flexible region of the linker but is itself, a very inflexible amino acid, so could play an important role in linker dynamics.

2.3 Cloning, Expression and Purification

2.3.1 Cloning of ^{WT}*Tma*DAH7PS

Prior to the beginning of this research, the gene for *Tma*DAH7PS (locus tag TM0343) had been amplified from *T. maritima* MSB8 (DSM 3109) genomic DNA and ligated into the pT7-7 vector.[83] The pT7-7 vector was transferred into *E. coli* XL1-Blue cells and the resulting construct transformed into *E. coli* BL21*(DE3) cell lines for protein expression. Sequencing analysis confirmed the presence of the correct sequence, within the expression plasmid.

2.3.2 Expression and Purification of ^{WT}*Tma*DAH7PS

The pT7-7 vector is under control of the *lac* promoter, which is induced by lactose. ^{WT}*Tma*DAH7PS expression was therefore induced by isopropyl-1-thiol-β-D-galactopyranoside (IPTG), which is a synthetic analogue of lactose that initiates

transcription of the lac operon but is not metabolised by the cell, as lactose is. Upon harvesting, the cell pellet was lysed and centrifuged, with the soluble fraction found to exhibit DAH7PS activity.

^{WT}*Tma*DAH7PS was purified from the crude lysate, using a purification procedure previously developed by Cross et al.[48, 82] As *T. maritima* is a hyperthermophile, the first step of the purification involved heat treatment of the cell lysate. When the cell lysate was heated to elevated temperatures, *E. coli* proteins present within the sample were precipitated out of solution, purifying ^{WT}*Tma*DAH7PS. In a slight modification to the procedure, heat debris was diluted with lysis buffer prior to centrifugation, improving collection of ^{WT}*Tma*DAH7PS following the heat treatment.

$(NH_4)_2SO_4$ (1 M) was added to the supernatant, which was loaded onto a Source^{TM} 15PHE column, pre-equilibrated with 1 M $(NH_4)_2SO_4$. As the concentration of $(NH_4)_2SO_4$ was lowered from 1 M to 0 M, in a step-wise fashion, ^{WT}*Tma*DAH7PS was eluted from the column. Following hydrophobic interaction chromatography (HIC), ^{WT}*Tma*DAH7PS was separated from remaining biomolecules, via size exclusion chromatography (SEC). HIC fractions corresponding to ^{WT}*Tma*DAH7PS were pooled and concentrated, before loading to the Superdex 10/300 SEC column. Fractions corresponding to purified ^{WT}*Tma*DAH7PS were pooled, concentrated and flash-frozen in liquid N_2, before storage at -80 °C.

Samples were taken at each stage of the purification procedure and analysed for increasing purity, by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) (Figure 2.4). Two litres of cell growth resulted in approximately 8 mg of purified ^{WT}*Tma*DAH7PS protein.

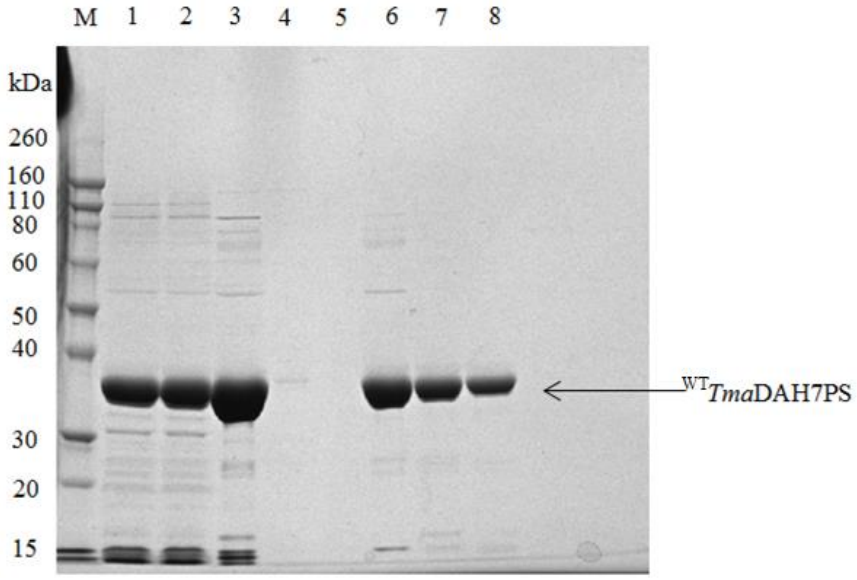

Figure 2.4: SDS-PAGE analysis of ^{WT}*Tma*DAH7PS purification. Lane M: molecular weight marker; Lane 1: supernatant after heat treatment; Lane 2: sample loaded to HIC column; Lane 3: pooled HIC fractions; Lane 4-8: fractions collected post SEC.

2.3.3 Site-directed Mutagenesis

^{V65P}*Tma*DAH7PS was previously constructed by site-directed mutagenesis and transformed into *E. coli* BL21*(DE3) cell lines, for protein expression. However, for ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS a QuikChange Lightning site-directed mutagenesis kit from Stratagene was utilised to introduce each mutation into the *Tma*DAH7PS pT7-7 plasmid, via extension of mutagenic primers.

The PCR product for each mutagenesis reaction was digested by DpnI and transformed into Stellar cells for protein expression. Colonies were grown at 37 °C overnight and the plasmids extracted for DNA sequencing. Sequencing results confirmed the presence of the correct sequence, within each expression plasmid.

2.3.4 Expression and Purification of *Tma*DAH7PS Linker Variants

The procedure outlined above for ^{WT}*Tma*DAH7PS, was utilised for expression and purification of ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS. For each linker variant, samples were taken at each stage of the purification procedure and analysed for increasing purity by SDS-PAGE. The increasing purity of ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS is presented in Figures 2.5-2.7 respectively. Two litres of cell growth resulted in approximately 7.5 mg of purified ^{V65P}*Tma*DAH7PS protein, 7 mg of ^{R64PP}*Tma*DAH7PS and 7 mg of ^{P68A}*Tma*DAH7PS.

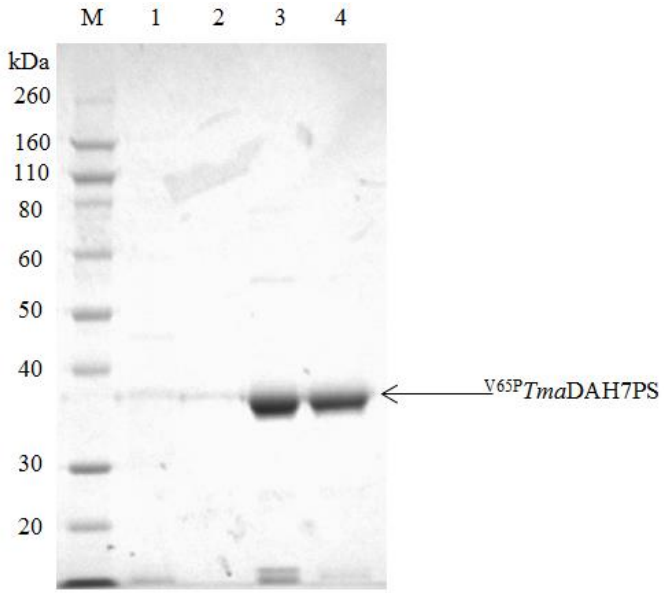

Figure 2.5: SDS-PAGE analysis of ^{V65P}*Tma*DAH7PS purification. Lane M: molecular weight marker; Lane 1: crude lysate; Lane 2: supernatant after heat treatment; Lane 3: pooled HIC fractions; Lane 4: pooled SEC fractions. Lanes are loaded at 1 mg/mL.

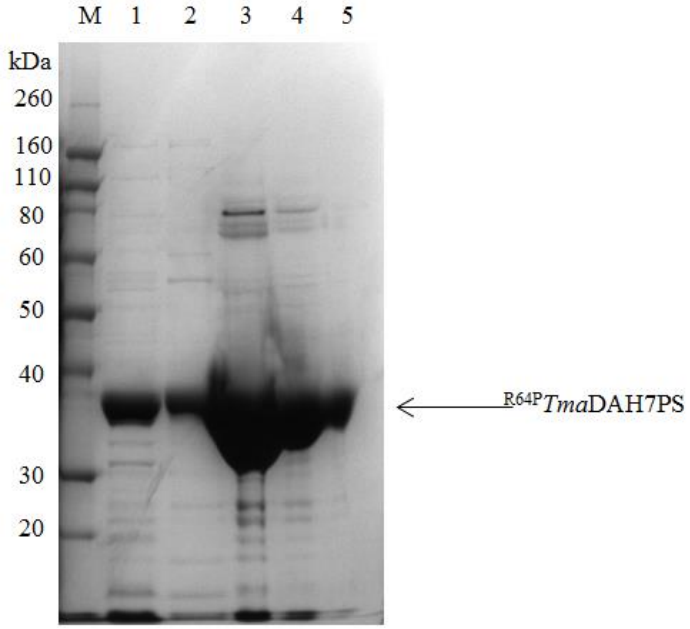

Figure 2.6: SDS-PAGE analysis of ^{R64P}*Tma*DAH7PS purification. Lane M: molecular weight marker; Lane 1: supernatant after heat treatment; Lane 2: pooled HIC fractions; Lane 3: sample loaded to SEC column; Lane 4: pooled SEC fractions; Lane5: concentrated ^{R64P}*Tma*DAH7PS sample.

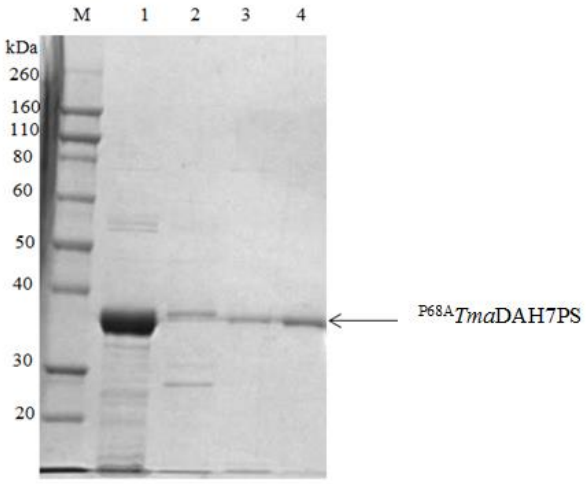

Figure 2.7: SDS-PAGE analysis of ^{P68A}*Tma*DAH7PS purification. Lane M: molecular weight marker; Lane 1: supernatant after heat treatment; Lane 2: pooled HIC fractions; Lane 3: pooled SEC fractions; Lane4: concentrated ^{P68A}*Tma*DAH7PS sample. Lanes are loaded at 1 mg/mL.

2.4 Molecular Weight Confirmation

The molecular mass of each protein was determined by electrospray ionisation mass spectrometry and compared to the theoretical mass, calculated from the amino acid sequence, with the ExPASy – ProtParam tool. ^{WT}*Tma*DAH7PS had a mass of 37,378.77 Da, which is nearly identical to the calculated mass of 37,378.16 Da. The mass of the linker variants ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS were also nearly identical to the theoretical masses; 37,376.87 Da, 37,319.65 Da and 37,352.74 Da compared to 37,376.15 Da, 37,319.09 Da and 37,352.13 Da respectively.

2.5 Kinetic Characterisation

A continuous spectrophotometric assay was performed at 60 °C for each enzyme, by measuring the disappearance of PEP absorbance at 232 nm. The assay was performed at 60 °C, rather than the optimal temperature of 80 °C for *Tma*DAH7PS activity,[78] due to the decreased stability of E4P at higher temperatures. The kinetic assay was initiated by the addition of E4P (300 µM) to the reaction mixture, which contained BTP buffer (50 mM BTP, pH 7.3), $MnSO_4$ (100 µM) PEP (200 µM) and enzyme.

2.5.1 Kinetic Parameters of ^{WT}*Tma*DAH7PS

For ^{WT}*Tma*DAH7PS, the apparent K_m values for PEP and E4P were 17 ±0.70 and 35 ±1.5 respectively (Figure 2.8). The k_{cat} was calculated to be 5.5 ±0.050 s^{-1}.

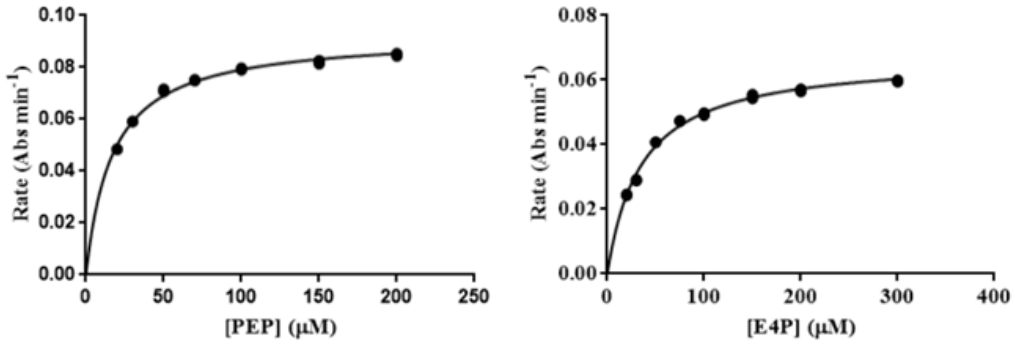

Figure 2.8: Michaelis-Menten curves of ^{WT}*Tma*DAH7PS. The K_m^{PEP} (left) was determined by fixing the [E4P] at 300 μM and varying the [PEP] from 20 to 200 μM. The K_m^{E4P} (right) was determined by fixing the [PEP] at 200 μM and varying the [E4P] from 20 to 300 μM. The assays were performed in triplicate and each data point is presented on the curve.

2.5.2 Kinetic Parameters of *Tma*DAH7PS Linker Variants

For ^{V65P}*Tma*DAH7PS, the apparent K_m values for PEP and E4P were 18 ±0.70 and 35 ±1.5 respectively (Figure 2.9). The k_{cat} was calculated to be 8.4 ±0.10 s^{-1}.

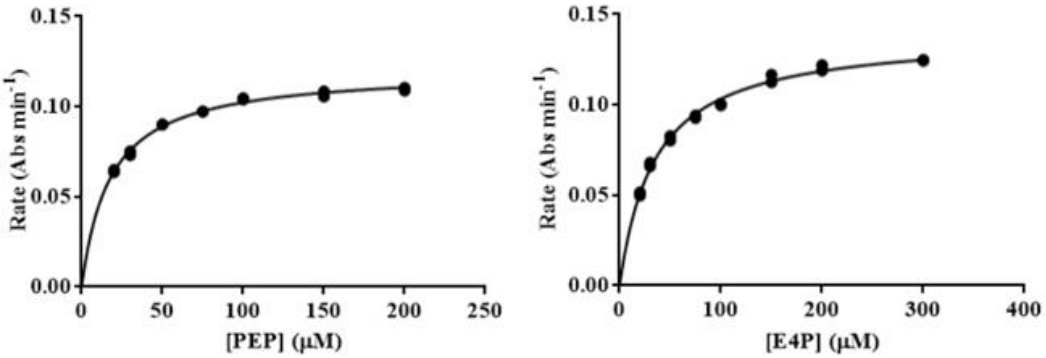

Figure 2.9: Michaelis-Menten curves of ^{V65P}*Tma*DAH7PS. The K_m^{PEP} (left) was determined by fixing the [E4P] at 300 μM and varying the [PEP] from 20 to 200 μM. The K_m^{E4P} (right) was determined by fixing the [PEP] at 200 μM and varying the [E4P] from 20 to 300 μM. The assays were performed in triplicate and each data point is presented on the curve.

For ^{R64P}*Tma*DAH7PS, the apparent K_m values for PEP and E4P were 19 ±1.0 and 21 ±1.5 respectively (Figure 2.10). The k_{cat} was calculated to be 5.5 ±0.060 s^{-1}.

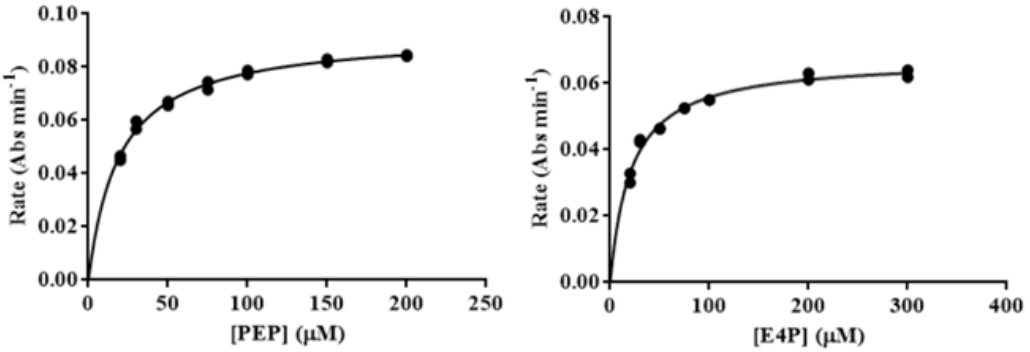

Figure 2.10: Michaelis-Menten curves of ^{R64P}*Tma*DAH7PS. The K_m^{PEP} (left) was determined by fixing the [E4P] at 300 µM and varying the [PEP] from 20 to 200 µM. The K_m^{E4P} (right) was determined by fixing the [PEP] at 200 µM and varying the [E4P] from 20 to 300 µM. The assays were performed in triplicate and each data point is presented on the curve.

For ^{P68A}*Tma*DAH7PS, the apparent K_m values for PEP and E4P were 18 ±1.0 and 18 ±1.0 respectively (Figure 2.11). The k_{cat} was calculated to be 8.1 ±0.10 s^{-1}.

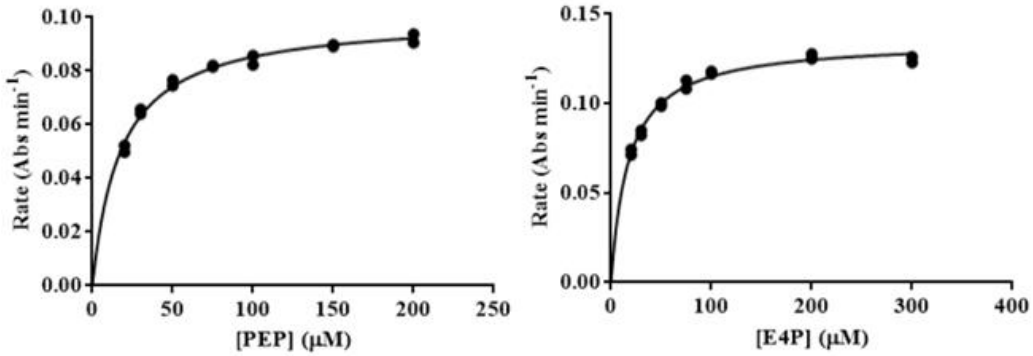

Figure 2.11: Michaelis-Menten curves of ^{P68A}*Tma*DAH7PS. The K_m^{PEP} (left) was determined by fixing the [E4P] at 300 µM and varying the [PEP] from 20 to 200 µM. The K_m^{E4P} (right) was determined by fixing the [PEP] at 200 µM and varying the [E4P] from 20 to 300 µM. The assays were performed in triplicate and each data point is presented on the curve.

The kinetic parameters for each of the *Tma*DAH7PS enzymes are presented in Table 2.1 for comparison.

Table 2.1: Kinetic parameters for *Tma*DAH7PS enzymes

Enzyme	K_m^{PEP} (μM)	K_m^{E4P} (μM)	k_{cat} (s^{-1})
^{WT}*Tma*DAH7PS	17 ±0.70	35 ±1.5	5.5 ±0.050
^{V65P}*Tma*DAH7PS	18 ±0.70	35 ±1.5	8.4 ±0.10
^{R64P}*Tma*DAH7PS	19 ±1.0	21 ±1.5	5.5 ±0.060
^{P68A}*Tma*DAH7PS	18 ±1.0	18 ±1.0	8.1 ±0.10

Overall, the apparent K_m values determined for ^{WT}*Tma*DAH7PS and the *Tma*DAH7PS linker variants are similar, and fit with the data previously reported for ^{WT}*Tma*DAH7PS,[82] indicating the concentration of each substrate required to achieve saturation of the enzyme, is not significantly altered by the linker mutations. Both ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS have increased k_{cat} values compared to ^{WT}*Tma*DAH7PS.

2.6 Feedback Inhibition Studies

The response of ^{WT}*Tma*DAH7PS and the *Tma*DAH7PS linker variants to increasing concentrations of either Tyr or Phe was determined by varying the ligand concentration from 0 μM to 400 μM. The assay outlined above was employed, with [E4P] fixed at 300 μM and [PEP] at 200 μM.

2.6.1 Inhibition of ^{WT}*Tma*DAH7PS

As expected, ^{WT}*Tma*DAH7PS is particularly sensitive to Tyr and was reduced to a constant 17.8% residual activity (Figure 2.12). This level of activity has been reported previously for ^{WT}*Tma*DAH7PS, at elevated concentrations of Tyr.[82]

^{WT}*Tma*DAH7PS is much less sensitive to Phe and even at the highest concentration measured, retains greater than 50% maximal activity.

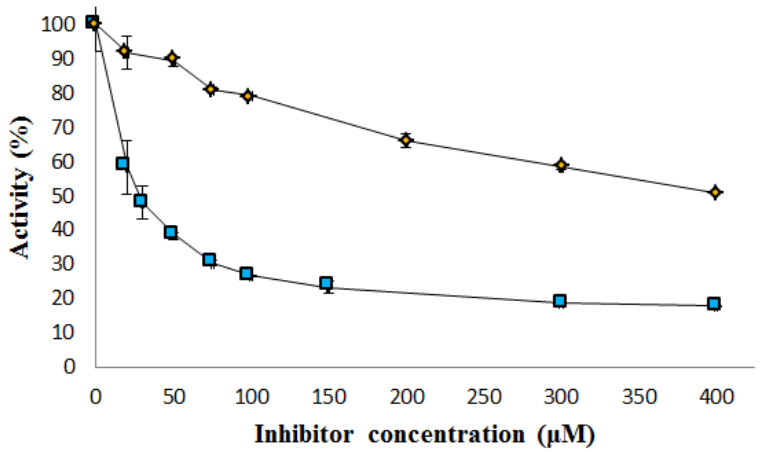

Figure 2.12: Response of ^{WT}*Tma*DAH7PS to increasing concentrations of Tyr (□) and Phe (◊). Each data point was measured in triplicate and the average value plotted.

2.6.2 Inhibition of *Tma*DAH7PS Linker Variants

^{V65P}*Tma*DAH7PS was less responsive to Tyr, compared to ^{WT}*Tma*DAH7PS and requires much higher concentrations to achieve the reduced level of inhibition observed (Figure 2.13). ^{V65P}*Tma*DAH7PS exhibits approximately 75% maximal activity, even in the presence of increasing concentrations of Tyr. Like ^{WT}*Tma*DAH7PS, ^{V65P}*Tma*DAH7PS is also less sensitive to Phe.

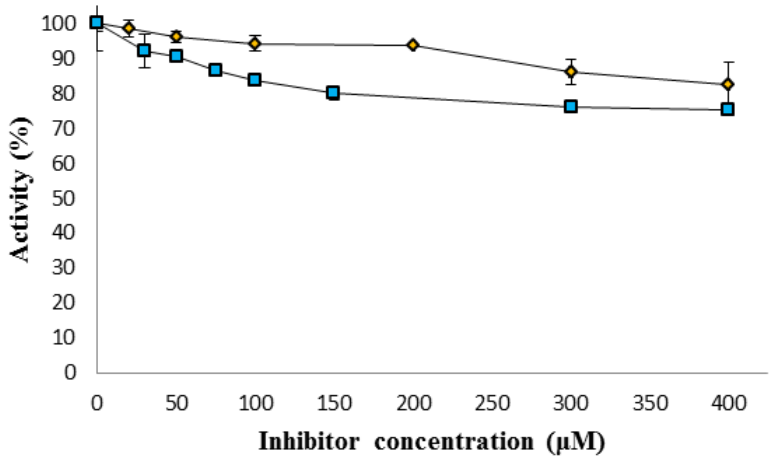

Figure 2.13: Response of ^{V65P}*Tma*DAH7PS to increasing concentrations of Tyr (□) and Phe (◊). Each data point was measured in triplicate and the average value plotted.

^{R64P}*Tma*DAH7PS exhibits similar activity to ^{WT}*Tma*DAH7PS in the presence of both Tyr and Phe (Figure 2.14). ^{R64P}*Tma*DAH7PS is slightly less sensitive to both Tyr and Phe, reduced to approximately 30% maximal activity in the presence of the highest concentration of Tyr.

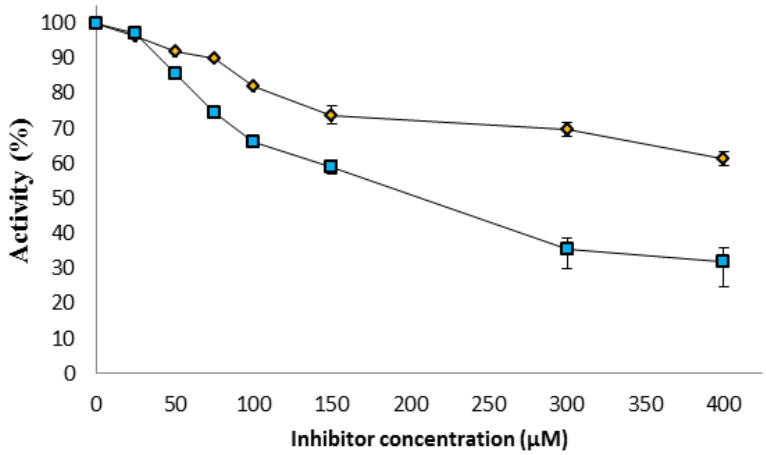

Figure 2.14: Response of ^{R64P}*Tma*DAH7PS to increasing concentrations of Tyr (□) and Phe (◊). Each data point was measured in triplicate and the average value plotted.

^{P68A}*Tma*DAH7PS was the most sensitive to Tyr, reduced to a residual activity less than 10% of the maximal rate (Figure 2.15). ^{P68A}*Tma*DAH7PS exhibited the greatest difference in inhibition for each ligand, reduced to approximately 70%

activity at the highest concentration of Phe. In comparison to ^{WT}*Tma*DAH7PS, ^{P68A}*Tma*DAH7PS is more sensitive to inhibition by Tyr but less sensitive to inhibition by Phe.

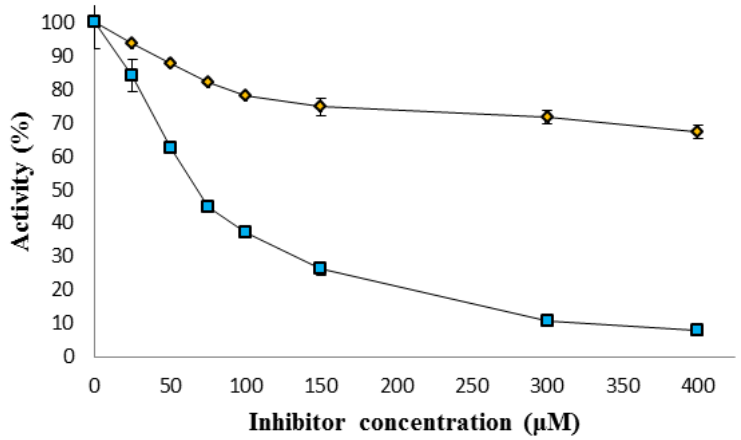

Figure 2.15: Response of ^{P68A}*Tma*DAH7PS to increasing concentrations of Tyr (□) and Phe (◊). Each data point was measured in triplicate and the average value plotted.

For clarity the response of the *Tma*DAH7PS enzymes to increasing concentrations of either Tyr (Figure 2.16) or Phe (Figure 2.17) is presented below.

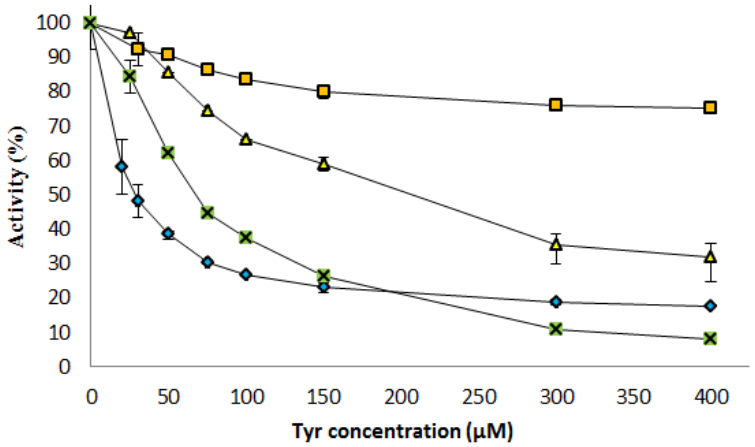

Figure 2.16: Response of ^{WT}*Tma*DAH7PS (◊), ^{V65P}*Tma*DAH7PS (□), ^{R64P}*Tma*DAH7PS (△) and ^{P68A}*Tma*DAH7PS (x) to increasing concentrations of Tyr. Each data point was measured in triplicate and the average value plotted.

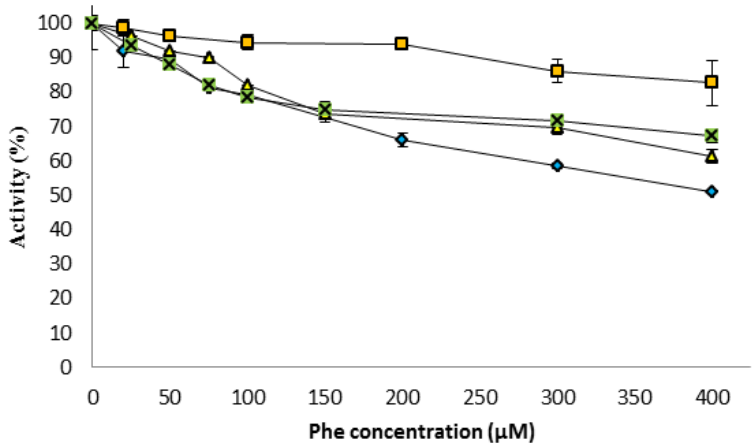

Figure 2.17: Response of ^{WT}*Tma*DAH7PS (◊), ^{V65P}*Tma*DAH7PS (□), ^{R64P}*Tma*DAH7PS (△) and ^{P68A}*Tma*DAH7PS (x) to increasing concentrations of Phe. Each data point was measured in triplicate and the average value plotted.

2.7 Differential Scanning Fluorimetry

Differential scanning fluorimetry (DSF) enables the thermal stability of proteins to be assessed, by measuring the temperature at which a protein becomes disordered.[84] DSF utilises a fluorescent dye that binds to the hydrophobic amino acids of a protein. Fluorescence is detected as the protein denatures and the hydrophobic amino acids become exposed, with increasing temperature.[84] DSF was performed to investigate the thermal stability of ^{WT}*Tma*DAH7PS and probe the impact of the linker mutations on thermal stability of the *Tma*DAH7PS linker variants.

2.7.1 Thermal Stability of ^{WT}*Tma*DAH7PS

For ^{WT}*Tma*DAH7PS, the melting temperature (T_m) was determined to be 93 °C in the absence of Tyr and 95 °C in the presence of Tyr (Figure 2.18). The melting temperatures are broadly in line with those previously reported for the first denaturation step of ^{WT}*Tma*DAH7PS.[82] ^{WT}*Tma*DAH7PS was reported to denature in two steps, the first occurring around 96 °C and the second at 100 °C.[82] The temperature range able to be accessed via DSF was not high enough to enable

detection of the second denaturation step for ^{WT}*Tma*DAH7PS. Therefore, only relative comparisons for the first denaturation step were able to be made.

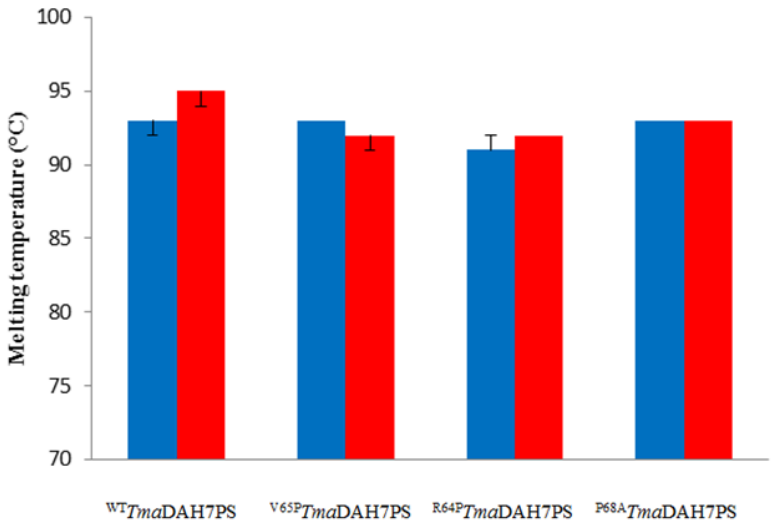

Figure 2.18: Thermal stability of ^{WT}*Tma*DAH7PS and the *Tma*DAH7PS linker variants, in the absence (blue) and presence (red) of Tyr (1 mM). Each data point was measured in triplicate and the average value plotted.

2.7.2 Thermal Stability of *Tma*DAH7PS Linker Variants

The *Tma*DAH7PS linker variants all exhibited similar melting temperatures to ^{WT}*Tma*DAH7PS, in both the absence and presence of Tyr (Figure 2.18). Similar melting temperatures suggest that the individual linker mutations have not significantly impacted the thermal stability of the *Tma*DAH7PS linker variants. Accurate determination of T_m by DSF, for the *Tma*DAH7PS enzymes, was limited by the temperature able to be achieved during the experiment. However, for investigating relative changes in thermal stability, the technique was suitable for preliminary analysis.

The T_m for each of the *Tma*DAH7PS enzymes, in the absence and presence of Tyr is presented in Table 2.2 for comparison.

Table 2.2: T_m for *Tma*DAH7PS enzymes in the absence and presence of 1 mM Tyr

Enzyme	Tyrosine	T_m (°C)
^{WT}*Tma*DAH7PS	-	93
	+	95
^{V65P}*Tma*DAH7PS	-	93
	+	92
^{R64P}*Tma*DAH7PS	-	91
	+	92
^{P68A}*Tma*DAH7PS	-	93
	+	93

2.8 Small Angle X-ray Scattering

Small angle X-ray scattering (SAXS) was employed to assess the solution structure of ^{WT}*Tma*DAH7PS and the *Tma*DAH7PS linker variants in the absence and presence of 1 mM Tyr. Each enzyme was analysed by SEC-SAXS, to collect scattering data from the monodisperse *Tma*DAH7PS species. Each enzyme was loaded to the column at a concentration of 10 mg/mL. Expression and purification of ^{R64P}*Tma*DAH7PS occurred after the scheduled SAXS beamline time, so this data was unable to be analysed within the time frame of this thesis.

2.8.1 Characterisation of ^{WT}*Tma*DAH7PS by Small Angle X-ray Scattering

Guinier analysis of the scattering profile for ^{WT}*Tma*DAH7PS in the absence of Tyr produced a radius of gyration (R_g) of 34.2 ±0.1 Å. The pair-distribution function (P(r)) was determined by indirect Fourier transformation and calculated the R_g to be 33.8 ±0.1 Å, and the maximum scattering particle (D_{max}) to be 131 Å.

The Porod volume (V_p), defined by the Porod invariant was calculated to be 220000 Å.[85]

Overlay of the SAXS data for ^{WT}*Tma*DAH7PS in the absence and presence of Tyr, revealed a noticeable change in the scattering profiles (Figure 2.19). In the presence of Tyr a radius of gyration (R_g) of 33.0 ±0.1 Å was determined by Guinier analysis, while the P(r) function calculated a R_g of 32.5 ±0.1 Å. The D_{max} was 103 Å and V_p was 215000 Å.[85] The scattering data is consistent with a conformational change for ^{WT}*Tma*DAH7PS in the presence of Tyr.

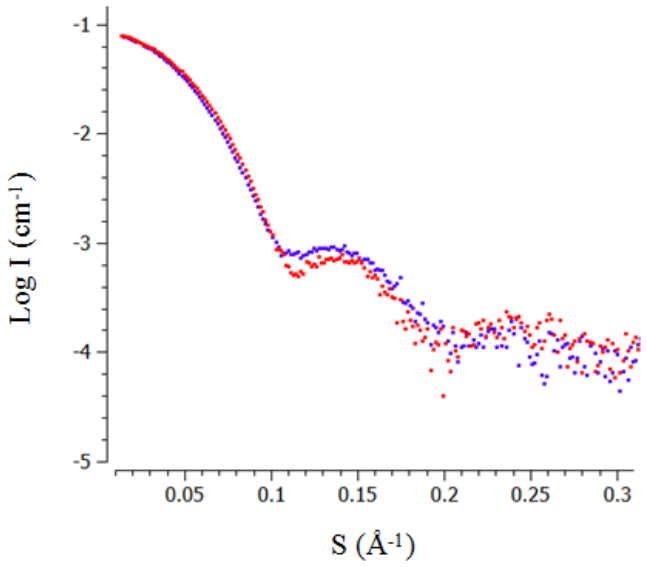

Figure 2.19: Overlay of the SAXS scattering profiles for ^{WT}*Tma*DAH7PS, in the absence (blue) and presence (red) of 1 mM Tyr.

2.8.2 Characterisation of *Tma*DAH7PS Linker Variants by Small Angle X-ray Scattering

Guinier analysis of the scattering profile for ^{V65P}*Tma*DAH7PS in the absence of Tyr produced a radius of gyration (R_g) of 33.6 ±0.3 Å, while the R_g determined by the P(r) function was 33.1 ±0.1 Å. The D_{max} was 115 Å and the V_p was calculated to be 223000 Å.[85]

Overlay of the SAXS data for ^{V65P}*Tma*DAH7PS in the absence and presence of Tyr, revealed minor changes in the scattering profiles (Figure 2.20). In the presence of Tyr a radius of gyration (R_g) of 33.2 ±0.1 Å was determined by Guinier analysis, while the P(r) function calculated a R_g of 32.5 ±0.1 Å. The D_{max} was 100 Å and V_p was 221000 Å.[85] The scattering data for ^{V65P}*Tma*DAH7PS suggests little conformational change in the presence of Tyr.

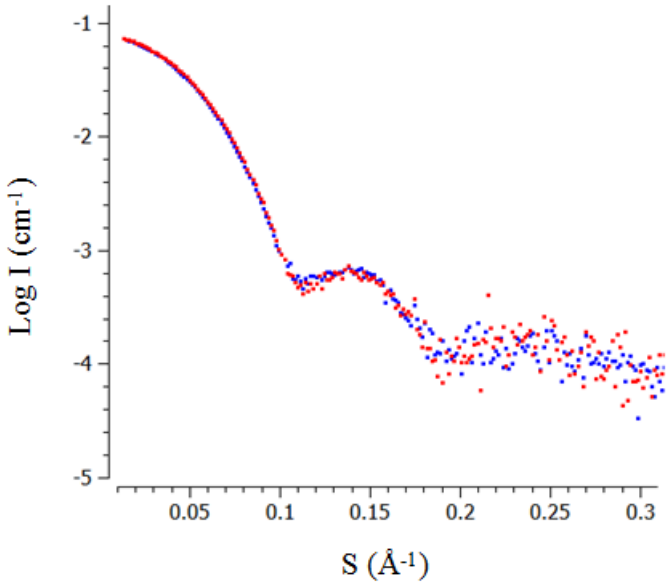

Figure 2.20: Overlay of the SAXS scattering profiles for ^{V65P}*Tma*DAH7PS, in the absence (blue) and presence (red) of 1 mM Tyr.

Guinier analysis of the scattering profile for ^{P68A}*Tma*DAH7PS in the absence of Tyr produced a radius of gyration (R_g) of 34.6 ±0.4 Å, while the R_g determined by the P(r) function was 34.2 ±0.1 Å. The D_{max} was 106 Å and the V_p was calculated to be 223000 Å.[85]

Overlay of the SAXS data for ^{P68A}*Tma*DAH7PS in the absence and presence of Tyr, revealed a large change in the scattering profiles (Figure 2.21). In the presence of Tyr a radius of gyration (R_g) of 32.8 ±0.2 Å was determined by Guinier analysis, while the P(r) function calculated a R_g of 32.2 ±0.1 Å. The D_{max} was 99 Å and V_p was 219000 Å.[85] The scattering data for ^{P68A}*Tma*DAH7PS suggests a large conformational change in the presence of Tyr. The

conformational change observed is more significant than that observed for ^{WT}*Tma*DAH7PS.

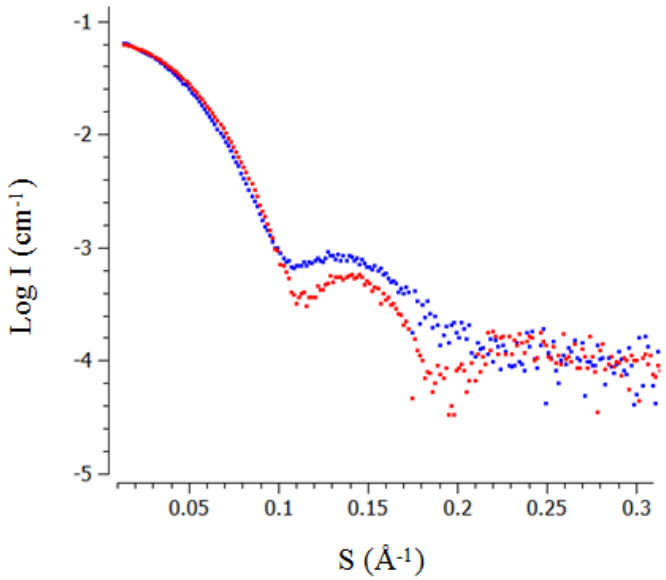

Figure 2.21: Overlay of the SAXS scattering profiles for ^{P68A}*Tma*DAH7PS, in the absence (blue) and presence (red) of 1 mM Tyr.

2.8.3 CRYSOL Fitting of SAXS Profiles for ^{WT}*Tma*DAH7PS

The change in protein conformation for ^{WT}*Tma*DAH7PS in the presence of Tyr can be more clearly detected when compared to the X-ray crystal structure. CRYSOL[86] was employed to predict the theoretical scattering profile for ^{WT}*Tma*DAH7PS from the X-ray crystal structure, and calculate the fit to the experimental SAXS profiles. For ^{WT}*Tma*DAH7PS, the fit to the open structure (PDB 1RZM) was better than the fit to the closed structure (PDB 3PG9) in the absence of Tyr, with χ^2 values of 0.26 and 1.90 respectively (figure 2.22). The superior fit to the open structure, suggests in the absence of ligand, aqueous ^{WT}*Tma*DAH7PS adopts a similar average conformation to that observed in the crystal structure.

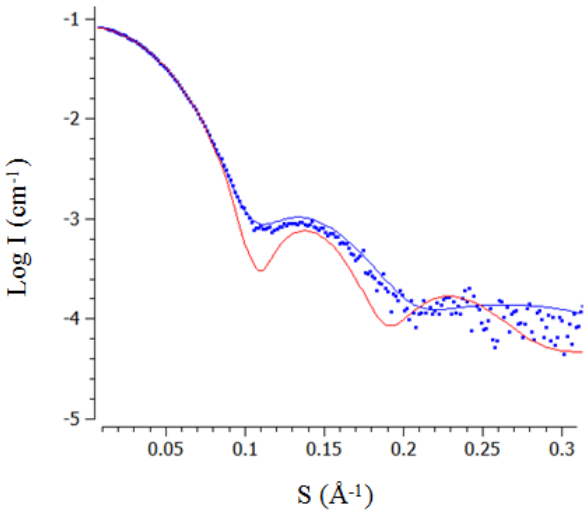

Figure 2.22: CRYSOL fit for ^{WT}*Tma*DAH7PS SAXS profile in the absence of Tyr (blue rectangles), to the theoretical scattering profiles calculated from the open crystal structure (blue line) and the closed crystal structure (red line).

For ^{WT}*Tma*DAH7PS, the fit to the closed structure was better than the fit to the open structure in the presence of Tyr, with χ^2 values of 0.62 and 1.93 respectively (figure 2.23). The superior fit to the closed structure suggests in the presence of ligand, aqueous ^{WT}*Tma*DAH7PS adopts a similar conformation.

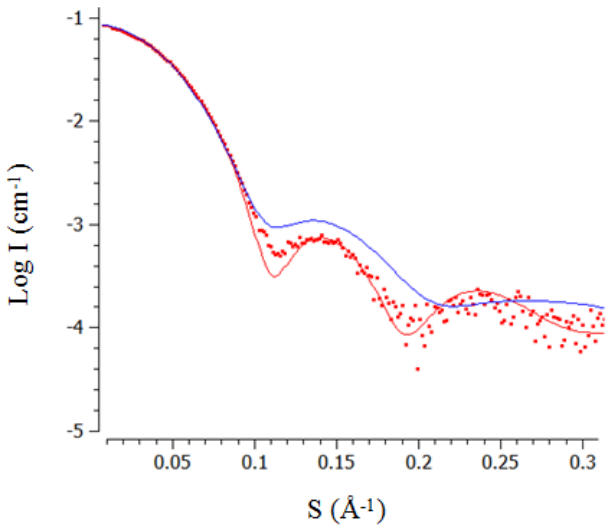

Figure 2.23: CRYSOL fit for ^{WT}*Tma*DAH7PS SAXS profile in the presence of Tyr (red rectangles), to the theoretical scattering profiles calculated from the open crystal structure (blue line) and the closed crystal structure (red line).

2.8.4 CRYSOL Fitting of SAXS Profiles for *Tma*DAH7PS Linker Variants

For ^{V65P}*Tma*DAH7PS, the fit to the open structure was worse than the fit to the closed structure in the absence of Tyr, with χ^2 values of 1.04 and 0.69 respectively (figure 2.24). This is in direct contrast to ^{WT}*Tma*DAH7PS and suggests that in the absence of ligand, aqueous ^{V65P}*Tma*DAH7PS adopts an altered conformation that more closely resembles the closed structure.

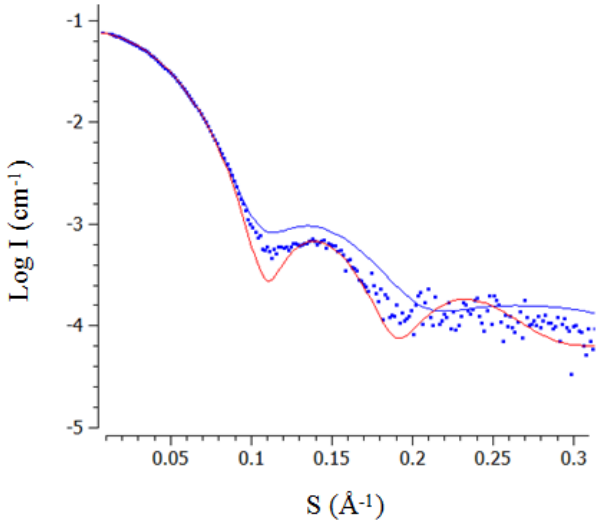

Figure 2.24: CRYSOL fit for ^{V65P}*Tma*DAH7PS SAXS profile in the absence of Tyr (blue rectangles), to the theoretical scattering profiles calculated from the open crystal structure (blue line) and the closed crystal structure (red line).

For ^{V65P}*Tma*DAH7PS, the fit to the closed structure was also better than the fit to the open structure in the presence of Tyr, with χ^2 values of 0.34 and 1.21 respectively (figure 2.25). The superior fit to the closed structure, suggests in the presence of ligand, aqueous ^{V65P}*Tma*DAH7PS adopts a similar conformation. Overall the conformation adopted by ^{V65P}*Tma*DAH7PS, more closely resembles the closed structure in both the absence and presence of Tyr. CRYSOL fitting indicates little change in protein conformation for ^{V65P}*Tma*DAH7PS in the presence of Tyr.

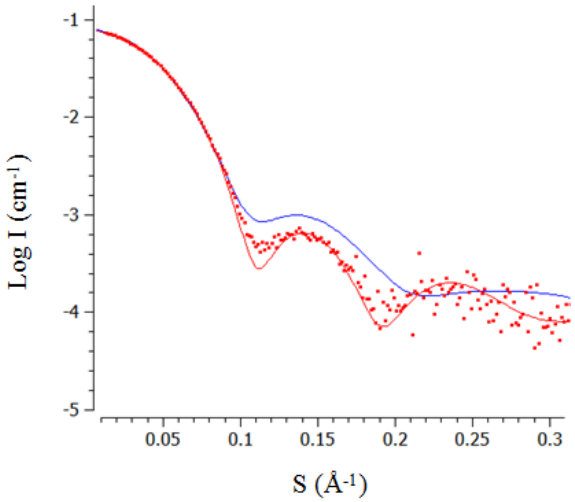

Figure 2.25: CRYSOL fit for ^{V65P}*Tma*DAH7PS SAXS profile in the presence of Tyr (red rectangles), to the theoretical scattering profiles calculated from the open crystal structure (blue line) and the closed crystal structure (red line).

For ^{P68A}*Tma*DAH7PS, the fit to the open structure was better than the fit to the closed structure in the absence of Tyr, with χ^2 values of 0.13 and 1.56 respectively (figure 2.26). The superior fit to the open structure, suggests in the absence of ligand, aqueous ^{P68A}*Tma*DAH7PS adopts a similar conformation as the crystal structure.

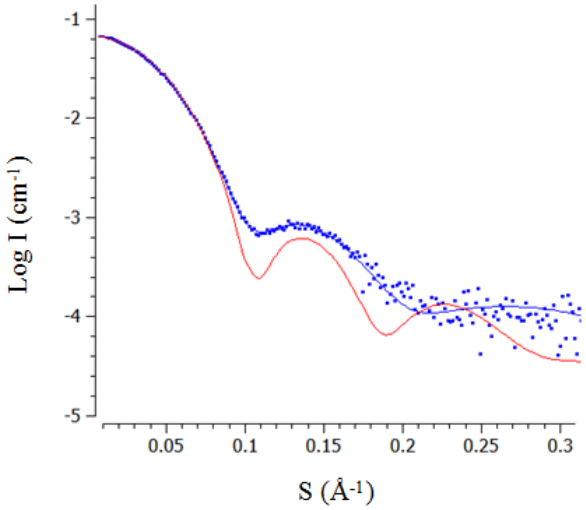

Figure 2.26: CRYSOL fit for ^{P68A}*Tma*DAH7PS SAXS profile in the absence of Tyr (blue rectangles), to the theoretical scattering profiles calculated from the open crystal structure (blue line) and the closed crystal structure (red line).

For ^{P68A}*Tma*DAH7PS, the fit to the closed structure was better than the fit to the open structure in the presence of Tyr, with χ^2 values of 0.25 and 2.01 respectively (figure 2.27). The superior fit to the closed structure, suggests in the presence of ligand, aqueous ^{P68A}*Tma*DAH7PS adopts a similar conformation. CRYSOL fitting confirms a large change in protein conformation for ^{P68A}*Tma*DAH7PS in the presence of Tyr. The change in conformation is more pronounced for ^{P68A}*Tma*DAH7PS compared to ^{WT}*Tma*DAH7PS.

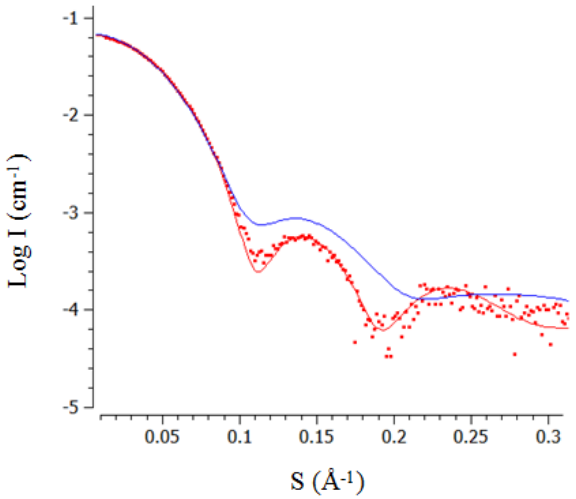

Figure 2.27: CRYSOL fit for ^{P68A}*Tma*DAH7PS SAXS profile in the presence of Tyr (red rectangles), to the theoretical scattering profiles calculated from the open crystal structure (blue line) and the closed crystal structure (red line).

2.9 Discussion

The characterisation of ^{WT}*Tma*DAH7PS provides a comparison for the *Tma*DAH7PS linker variants, supporting the conclusions drawn from each result.^{WT}*Tma*DAH7PS was successfully purified using an established procedure and found to have kinetic activity similar to that previously reported.[82] ^{WT}*Tma*DAH7PS is responsive to Tyr and Phe, with more potent inhibition achieved by Tyr. ^{WT}*Tma*DAH7PS has high thermal stability in both the absence and presence of Tyr.

The allosteric regulation mechanism for ^{WT}*Tma*DAH7PS, involves a large conformational change in the presence of Tyr. The scattering data is consistent with conformational change for ^{WT}*Tma*DAH7PS, with decreased R_g, D_{max} and V_p when Tyr is present. The conformational change was observed via overlay of the SAXS data, in the absence and presence of Tyr, which confirmed differences in the overall SAXS profiles. The conformational change for ^{WT}*Tma*DAH7PS was more readily observed when comparing the SAXS data to the X-ray crystal structures, with ^{WT}*Tma*DAH7PS adopting a similar average conformation to either the open crystal structure, in the absence of Tyr, or the closed crystal structure, when Tyr was present.

The *Tma*DAH7PS linker variants were purified in the same manner as ^{WT}*Tma*DAH7PS, with good success. Melting temperatures for the *Tma*DAH7PS linker variants were almost identical to ^{WT}*Tma*DAH7PS, indicating the thermal stability of the enzymes had not been compromised by the individual linker mutations. Overall the kinetic properties amongst each enzyme were similar. Of interest were the increased k_{cat} values for ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS, which were calculated to be 8.4 ±0.10 s^{-1} and 8.1 ±0.10 s^{-1} respectively. Each value is over 1.5 times greater than that of ^{WT}*Tma*DAH7PS (5.5 ±0.050 s^{-1}), suggesting altered flexibility of the linker region can impact enzyme activity. However, these results should not be over-interpreted, as ^{WT}*Tma*DAH7PS displays some variability in k_{cat}, evidenced by slight differences in the reported value.[82] Nevertheless the increase in activity observed could be due to decreased sampling of the fully closed conformation. For ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS, altered flexibility of the linker may enable adoption of altered angles that favour alternate conformations and reduce time spent sampling the closed conformation.

The reduced sensitivity of ^{V65P}*Tma*DAH7PS to Tyr, suggests that reduction in flexibility has disrupted linker dynamics, impacting the ability of the enzyme to adopt the closed conformation in the presence of Tyr. Higher concentrations of Tyr still reduce enzyme activity, so the closed conformation, or a conformation

similar, may well be accessible. However, decreased flexibility may have reduced the level of stabilisation Tyr would normally provide, leading to a decrease in sampling of the closed conformation and a reduction in inhibition by Tyr. This result implies that either the relative stability of the open vs closed conformation has been altered, or that the closed conformation ^{V65P}*Tma*DAH7PS can access, is not the same conformation that ^{WT}*Tma*DAH7PS can access.

What is clear is the impact the reduction of linker flexibility has on allosteric regulation of ^{V65P}*Tma*DAH7PS. Fitting to the crystal structure models, suggests ^{V65P}*Tma*DAH7PS adopts a similar conformation to the closed wild-type protein structure, in both the absence and presence of Tyr. However, the adopted conformation does not seem to enable full inhibition by Tyr.

^{R64P}*Tma*DAH7PS exhibits the most similar characteristics to ^{WT}*Tma*DAH7PS but is slightly less sensitive to Tyr and Phe. Potentially the reduced flexibility of position 64 is not as important as flexibility at position 65 for linker dynamics. The reduced linker flexibility of ^{R64P}*Tma*DAH7PS has impaired the allosteric response somewhat but has had little effect on kinetic activity. Although SAXS analysis was unable to be performed for ^{R64P}*Tma*DAH7PS, the similarity to ^{WT}*Tma*DAH7PS enables prediction of gross conformational change in the presence of Tyr. However, the extent to which the fully closed conformation is adopted by ^{R64P}*Tma*DAH7PS, may be somewhat reduced compared to ^{WT}*Tma*DAH7PS.

^{P68A}*Tma*DAH7PS has improved sensitivity to Tyr. Like ^{WT}*Tma*DAH7PS, the allosteric regulation mechanism for ^{P68A}*Tma*DAH7PS involves a large conformational change, in the presence of Tyr. The scattering data is consistent with a change in conformation for ^{P68A}*Tma*DAH7PS, with overlay of the SAXS data in the absence and presence of Tyr, highlighting significant differences in the overall SAXS profiles. CRYSOL fitting suggests ^{P68A}*Tma*DAH7PS adopts a very similar conformation to the open and closed structures of the wild-type protein in

response to Tyr binding. Increased flexibility of the ^{P68A}*Tma*DAH7PS linker has improved enzyme activity and the potency of the allosteric response.

Biochemical characterisation and comparison of ^{WT}*Tma*DAH7PS and the *Tma*DAH7PS linker variants, confirms the role of the linker region in allosteric regulation of *Tma*DAH7PS. The flexibility of the linker region appears to be integral to the mechanism of allostery employed by *Tma*DAH7PS, and alteration to linker flexibility, can significantly impact allosteric regulation of the enzyme.

Chapter Three

Protein Dynamics in Allosteric Regulation of *Thermotoga maritima* DAH7PS

3.1 Introduction

Gross conformational change and alterations in conformational flexibility have been associated with allostery and are defined as dynamic processes. For a protein, dynamic processes occur over a range of timescales, including very fast fluctuations of atoms, which take place on the picosecond timescale, to conformational rearrangement on the millisecond timescale.[71,87] Over the last decade it has been realised that the molecular motions conferring allostery, take place with and without ligand binding, leading to proteins being described as a conformational ensemble of dynamic sub-states.[71,87]

For *Tma*DAH7PS, the mechanism of allostery involves gross conformational change that may arise from alteration in protein dynamics, upon ligand binding. *In silico* techniques were therefore valuable to study in detail, changes in protein dynamics, occurring within the enzyme structure with Tyr binding. The modelling techniques employed to study molecular motion of *Tma*DAH7PS include molecular dynamics (MD) and targeted molecular dynamics (TMD), conducted with the software NAMD.[88] The MD and TMD simulation work is presented in this chapter and supports the role of linker dynamics in the mechanism of allostery for *Tma*DAH7PS.

3.2 Protein Dynamics

Proteins are inherently dynamic biomolecules, which are able to sample an ensemble of conformations. [71, 89] Present within the protein ensemble are a number of pre-existing sub-states that resemble either the protein ground state, or in some cases a discrete conformer higher in energy.[90] The occupancy of a particular sub-state is dependent on the relative energy of the conformation the sub-state represents.[66] Conformations lower in energy are sampled more often and are the predominant species in solution.[89-90] The model of conformational selection is based upon the energy landscape theory, where different protein conformations exist in dynamic equilibrium.[89-91] Each protein conformation is separated by an energy barrier and it is the relative energies of the conformations and their accessibility, as determined by the height of the energy barriers that separate them, that influences protein distribution over the free energy landscape.[89,91]

Ligand binding to a protein alters the relative energy of sub-states in the conformational ensemble. In addition, the frequency of sub-state sampling may be altered by ligand binding, shifting the energy profile of a protein and allowing different protein conformations to be accessed (Figure 3.1).[67,89] A ligand might bind to only a few specific conformations and stabilise them. However, stabilisation modifies the entire ensemble and consequently the population of each sub-state.[66] The binding interaction between a ligand and higher energy sub-state causes a population shift and leads to a change in observed conformer. The species that are occupied, and the time it takes for conformational change to occur, depend on the individual energy profile of a protein.[89]

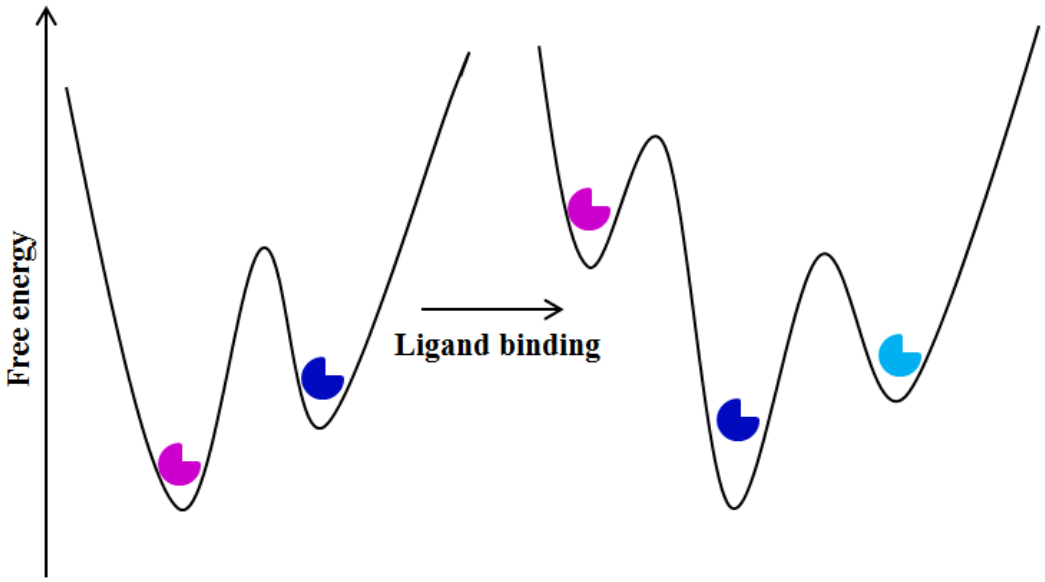

Figure 3.1: Free energy landscape of a protein, with the protein ground state represented in purple and discrete protein conformers in blue. Similar protein conformations are coloured alike.

The number of studies reporting a correlation between a protein's equilibrium dynamics and structural changes involved in protein function, has increased in recent years, supporting the idea that proteins have an intrinsic ability to sample functional conformations that approximate those required for biological activity.[92-93] The capacity of proteins to sample open and closed forms in the absence of ligand has been described, where the presence of ligand results in stabilisation of one form (usually closed), leading to a shift in the protein ensemble.[92] The concept of "population shift" for a protein ensemble is the foundation for current understanding of allostery, determined by alteration in enthalpy (gross conformational change) and entropy (changes in protein dynamics and conformational flexibility).[67, 72, 93]

3.2.1 *Tma*DAH7PS Protein Dynamics

The transition from the open to closed conformation of *Tma*DAH7PS is facilitated by protein dynamics. For *Tma*DAH7PS it is predicted that the open conformation

represents the protein ground state, and is occupied within the protein ensemble the majority of the time. The inactive conformer, observed upon Tyr binding, is likely higher in energy, so not occupied as frequently as the active, Tyr free conformer.

When Tyr binds, the relative energy of individual sub-states within the protein ensemble changes, enabling the transition from the open to closed conformation.[94] The change in energy landscape in response to Tyr binding, results in a large average conformational adjustment for *Tma*DAH7PS.[46] It seems likely that a dynamic equilibrium between the open and closed forms of *Tma*DAH7PS exists, and that the presence of Tyr alters the position of the equilibrium, favouring the closed enzyme conformation.[46]

The dynamic nature of *Tma*DAH7PS is pivotal to the mechanism of allosteric regulation exhibited by the enzyme.[48] The transition from the open to closed conformation of *Tma*DAH7PS is facilitated by the flexible linker region, which enables the regulatory domains to re-position upon Tyr binding. The function of the linker region is likely not only for flexibility but may also contribute to an allosteric propagation pathway; where subtle changes in protein dynamics enable the closed conformation to be accessed, upon Tyr binding.[56,76] A propagation pathway may be pre-encoded by the amino acids of the linker, so that successive states are encoded by the previous one, driving the conformational change of *Tma*DAH7PS in the presence of Tyr.[76]

3.3 Molecular Dynamics

Molecular dynamics (MD) is a computational simulation of the movement of atoms in a molecule over time. The time dependent behaviour of a molecular system is calculated by applying Newton's equation of motion to the individual atoms of the molecular system.[88-95] In a MD simulation the Newtonian equation considers time and atomic positions and is of the form:

$$\frac{m_a \partial^2(r_a)}{\partial t^2} = -\frac{\partial U_{total}}{\partial r_a}$$

where m_a is the mass of atom a, $\frac{\partial^2(r_a)}{\partial t^2}$ is the acceleration of atom a; in terms of the positional vector (r_a) of atom a and the simulation time (t), and $-\frac{\partial U_{total}}{\partial r_a}$ is the force acting on atom a, calculated from the total potential energy (U_{total}) of every atomic position.[95] U_{total} is dependent on each atomic position and is represented by the MD force field, which considers bonding and non-bonding interactions for the system.[96]

For a MD simulation, periodic boundary conditions (PBC) are set to approximate an infinite system, via a unit cell. PBC places the molecular system being modelled into a unit cell that is considered to have infinite images in space, so that any atom exiting the unit cell is replaced by its image, entering the opposite side of the unit cell.[97]

In MD it is unpractical to calculate every long range interaction exactly and for MD simulations performed with NAMD, long range Van der Waals and electrostatic interactions are cut off at a distance of 25 Å.[88] To calculate long range interactions, the particle-mesh Ewald (PME) method is used in combination with the periodicity of PBC.[98-99] PME divides the total electrostatic interaction energy into a summation of short range (direct sum) and long range (Ewald sum) electrostatic interactions.[99]

3.3.1 Molecular Dynamics Simulations for *Tma*DAH7PS

MD simulations were conducted to examine the dynamic properties of the linker region for *Tma*DAH7PS. MD simulations were set up to survey the effect of reducing the flexibility of the linker for two systems: the open, ligand free system and the closed, Tyr bound system.

*Tma*DAH7PS crystal structures capturing both the open (PDB: 1RZM) and closed (PDB: 3PG9) conformation of the enzyme were used as starting points for each simulation. The tetrameric structure for each complex was generated from the associated dimer in the asymmetric unit, based on crystallographic symmetry. Preparation of each structure involved the removal of all ligands and replacement of active site Cd^{2+} ions with Mn^{2+} ions; as kinetic assays were performed in the presence of Mn^{2+}. Open and closed systems were solvated in an explicit TIP3 water box, which enabled movement of the regulatory domains; without interaction with periodic boundary images, during the MD simulations. Na^{+} and Cl^{-} ions were added to each system to neutralise the overall net charge, with a minimum 5 Å distance between individual ions and between the enzyme and each ion.

NAMD was used to carry out all MD simulations, using the CHARMM all atom parameter set 22.[88] MD simulations were run using 1024 cores of the IBM Blue Gene/P computer from the University of Canterbury HPC supercomputing and services facility. Electrostatic interactions were calculated using the PME method, with a Van der Waals cut-off of 12 Å. Each simulation was performed at 1 atmosphere and 363 K, the known optimal temperature for *T. maritima* activity. For each simulation an energy minimisation of 5000 steps, using the conjugate gradient energy minimisation method was run. Time steps of 2 fs were utilised and a trajectory written out every 100 ps.

3.3.2 Molecular Dynamics Simulations for the Open Conformation of ^{WT}*Tma*DAH7PS

Classical MD simulations for the open conformation of ^{WT}*Tma*DAH7PS were performed by Lang.[100] A 335.1 ns trajectory for the open conformation of ^{WT}*Tma*DAH7PS was kindly provided for analysis, enabling comparison to the ^{V65P}*Tma*DAH7PS trajectories presented here. The MD simulation for the open conformation of ^{WT}*Tma*DAH7PS, presents a relatively large increase in protein

backbone RMSD, with respect to the open crystal structure (Figure 3.2). Changes in RMSD occur rapidly, with RMSD greater than 8 Å relative to the starting structure, achieved within the first 50 ns of the simulation. For the remainder of the simulation the RMSD proceeds to oscillate between 8 Å and 10 Å, suggesting the protein undergoes a series of conformational rearrangements, without reaching an energy minimum that would suggest adoption of a stabilised conformer.

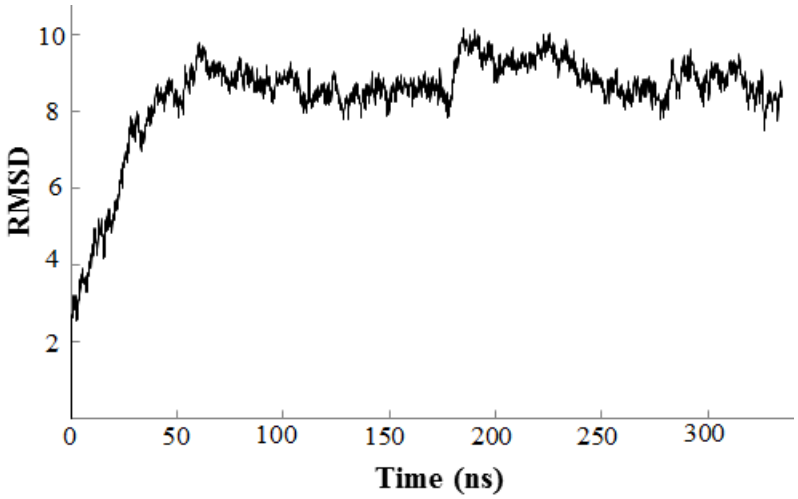

Figure 3.2: Backbone RMSD of ^{WT}*Tma*DAH7PS during a 335.1 ns MD simulation.

Visual examination of the MD trajectory for the open conformation of ^{WT}*Tma*DAH7PS confirmed conformational rearrangement of the regulatory domains throughout the simulation (Figure 3.3). In contrast, the core catalytic TIM barrels were relatively motionless, with limited conformational adjustment observed during the simulation. Conformational sampling was facilitated by the flexible inter-domain linker, which enabled the regulatory domains to explore a significant amount of the conformational space accessible to the protein. However, sampling of the fully closed conformation was unable to be observed via MD, which was predicted considering the high degree of freedom imparted by the flexible linker to the regulatory domains.

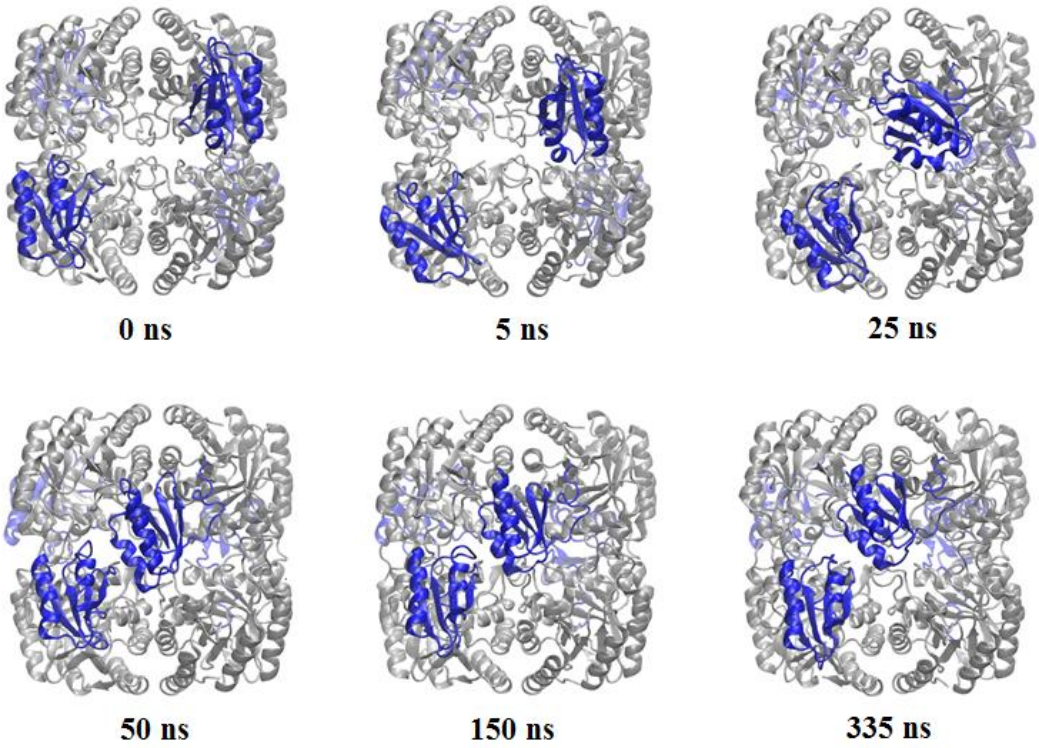

Figure 3.3: Conformational poses of ^{WT}*Tma*DAH7PS during a 335.1 ns MD simulation of the open conformation.

3.3.3 Molecular Dynamics Simulations for the Open Conformation of ^{V65P}*Tma*DAH7PS

The initial MD simulation for the open conformation of ^{V65P}*Tma*DAH7PS was run for 256.9 ns. Overall protein backbone RMSD values indicate relatively small deviations in backbone conformation, with respect to the open crystal structure (Figure 3.4). RMSD greater than 2 Å was achieved almost immediately upon initiating the simulation. However, as the simulation progressed the RMSD fluctuations were small, reaching a maximum of 3 Å before proceeding to oscillate around 2.5 Å.

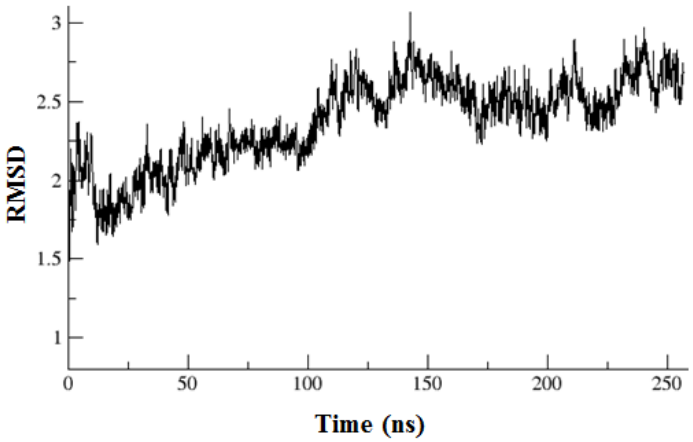

Figure 3.4: Backbone RMSD of ^{V65P}*Tma*DAH7PS during a 256.9 ns MD simulation.

As for ^{WT}*Tma*DAH7PS, visual examination of the MD trajectory for the open conformation of ^{V65P}*Tma*DAH7PS confirmed conformational rearrangement of the regulatory domains (Figure 3.5). However, the conformational adjustments were smaller for ^{V65P}*Tma*DAH7PS, which appeared to sample a more limited conformational space. During the simulation regulatory domains from diagonally opposite monomers were not observed to interact, nor undergo movement in the direction necessary to adopt the closed conformation. It seems likely that the reduced flexibility of the linker has impaired the ability of ^{V65P}*Tma*DAH7PS to adopt the closed conformation, or at least the range of conformations available to this modified protein.

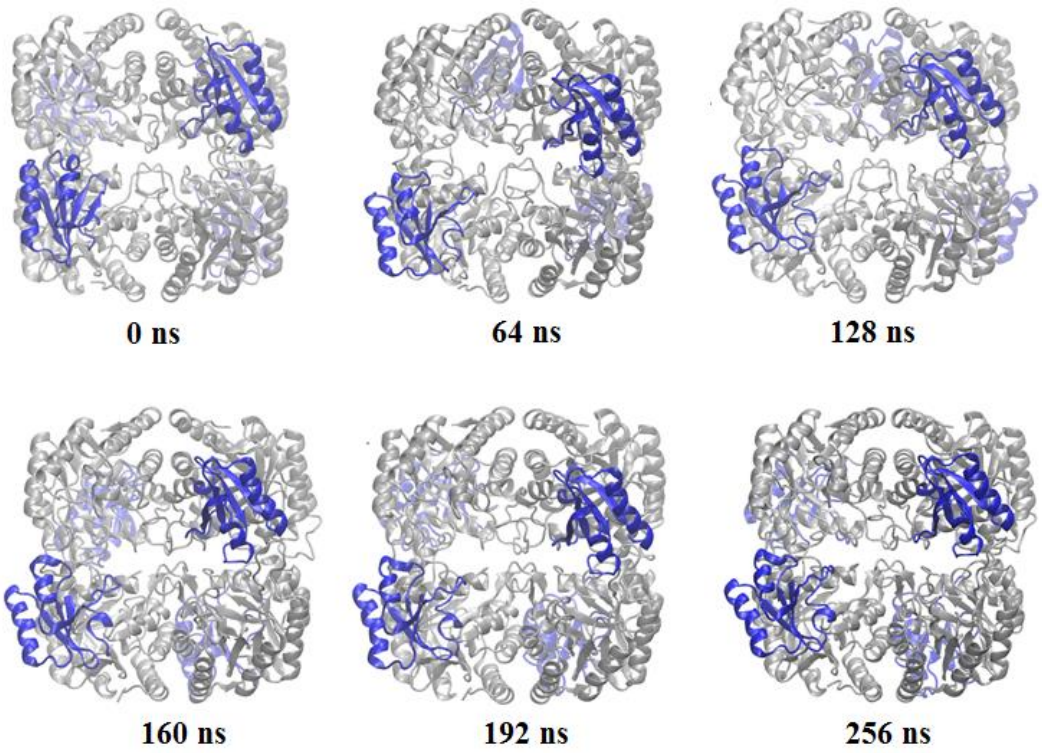

Figure 3.5: Conformational poses of ^{V65P}*Tma*DAH7PS during a 256.9 ns MD simulation of the open conformation.

Second and third classical MD simulations were run for the open conformation of ^{V65P}*Tma*DAH7PS, using the same parameters but different initial starting velocities to capture alternate conformational space explored by the enzyme. The MD simulations were run for 256.5 ns and 64 ns respectively. Protein backbone RMSD values indicate slightly larger deviations in backbone conformation, with respect to the open crystal structure (Figure 3.6-3.7). However, the RMSD oscillations were not as great as for the open conformation of ^{WT}*Tma*DAH7PS. As before, neither MD simulation revealed interaction between the regulatory domains or movement of the regulatory domains in the direction required to adopt the closed conformation (Figure 3.8).

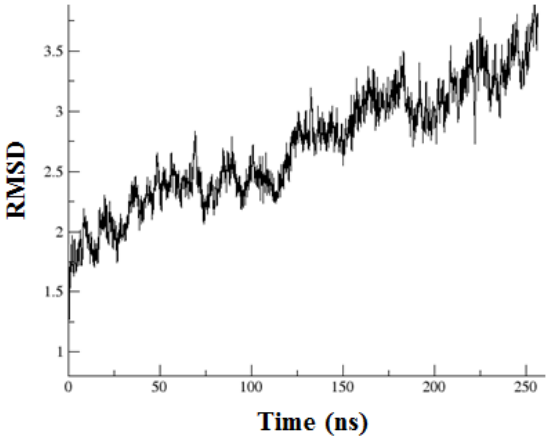

Figure 3.6: Backbone RMSD of ^{V65P}*Tma*DAH7PS during a 256.5 ns MD simulation.

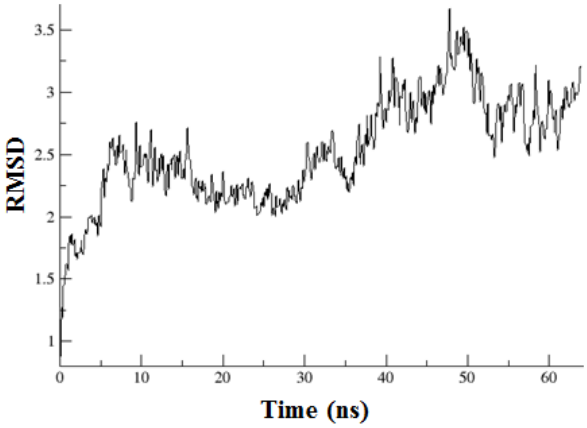

Figure 3.7: Backbone RMSD of ^{V65P}*Tma*DAH7PS during a 64 ns MD simulation.

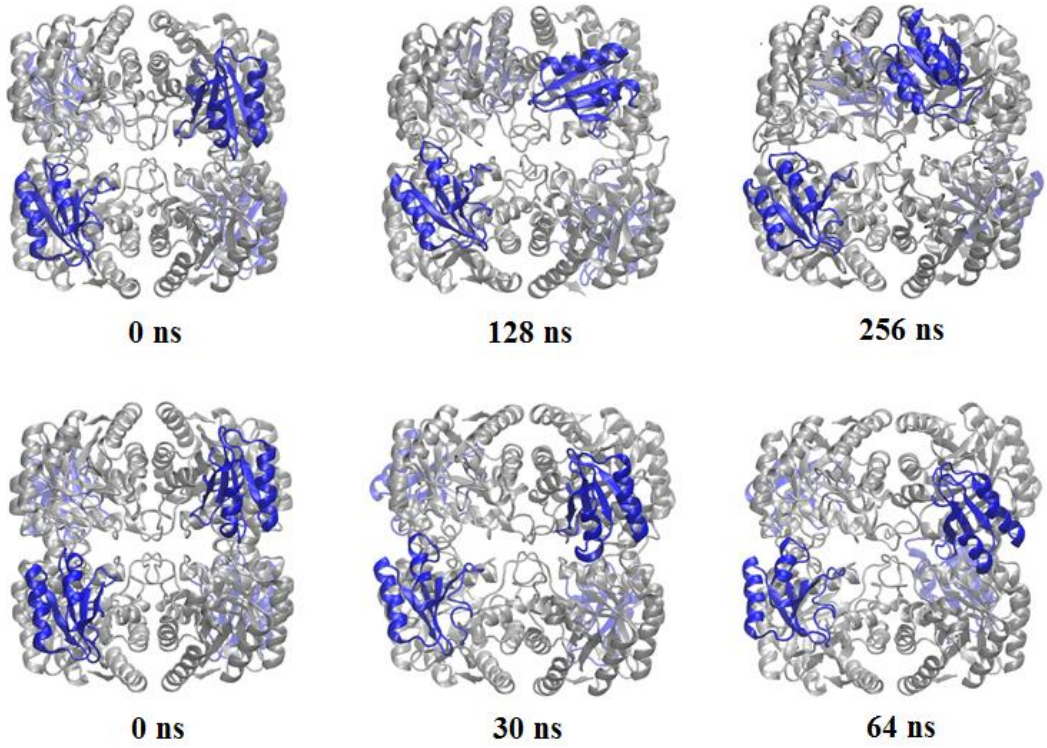

Figure 3.8: Conformational poses of ^{V65P}*Tma*DAH7PS during 256.5 ns (above) and 64 ns (below) MD simulations of the open conformation.

3.3.4 Molecular Dynamics Simulations for the Closed Conformation of ^{V65P}*Tma*DAH7PS

The closed conformation of ^{V65P}*Tma*DAH7PS was simulated for 101.4 ns without Tyr bound and 67.1 ns with Tyr bound. Protein backbone RMSD values indicate relatively small deviations in backbone conformation, with respect to the closed crystal structure in both the absence (Figure 3.9) and presence of Tyr (Figure 3.10). The RMSD increases slowly over the course of each simulation, oscillating between 1.5 Å and 2.5 Å.

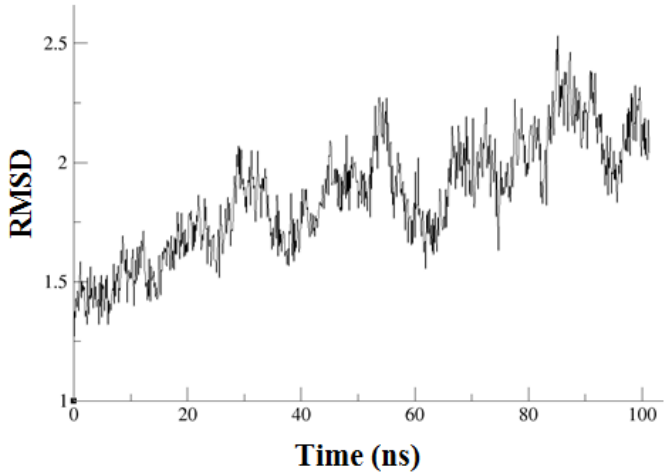

Figure 3.9: Backbone RMSD of ^{V65P}*Tma*DAH7PS during a 101.4 ns MD simulation, without Tyr bound.

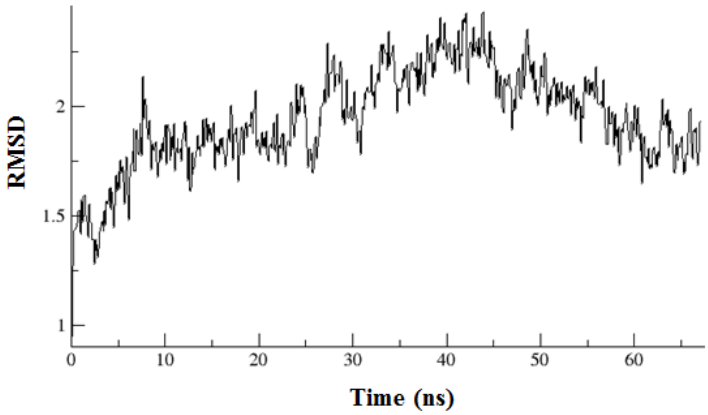

Figure 3.10: Backbone RMSD of ^{V65P}*Tma*DAH7PS during a 67.1 ns MD simulation, with Tyr bound.

Visual examination of the MD trajectory for each simulation revealed minimal displacement of the regulatory domains from starting atomic coordinates (Figure 3.11). The regulatory domains remained in the closed position for the entirety of the simulation. Even in the absence of Tyr the closed conformation was still adopted, suggesting the closed conformation is relatively stable compared to the open conformation of ^{V65P}*Tma*DAH7PS. Regulatory domain pairs above and below the tetramer appeared to interact normally with each other and with the Tyr

bound, indicating the linker substitution had little or no effect, on the ability of the enzyme to adopt the closed conformation and interact with Tyr.

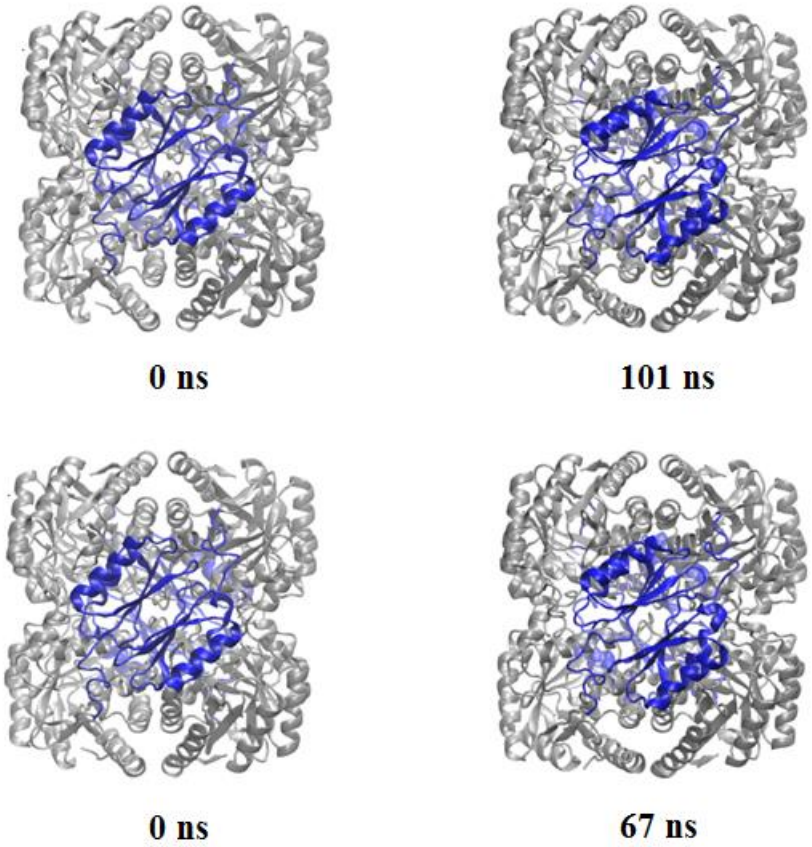

Figure 3.11: Conformational poses of ^{V65P}*Tma*DAH7PS during a 101.4 ns (above) and 67.1 ns (below) MD simulation of the closed conformation.

A frame by frame distance comparison between the closed conformation with and without Tyr bound was performed by measuring the distance between the α-carbons of His29 and Ser31, which are present in the pocket where Tyr binds and form a known hydrogen bond between adjacent regulatory domains (Figure 3.12). No significant difference in the distance between the residues was found. The average distance measured was 11.4 Å in the presence and absence of Tyr. This seems to indicate that the linker substitution had little or no effect on Tyr binding, as key interactions between residues present at the site of Tyr binding remained intact for ^{V65P}*Tma*DAH7PS.

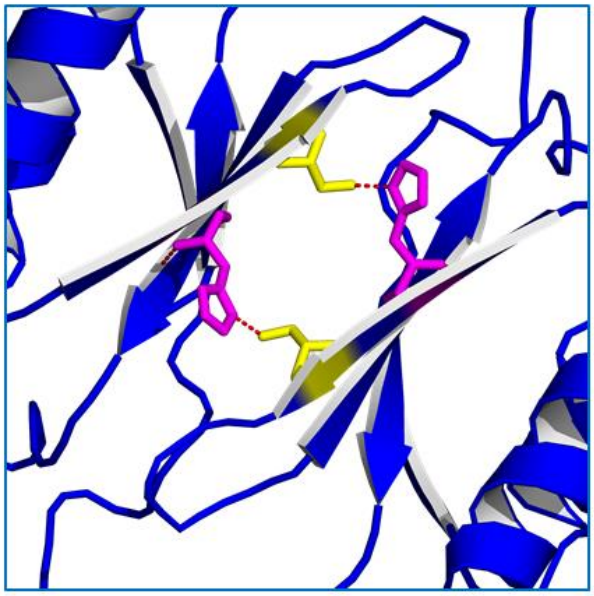

Figure 3.12: Cartoon representation of the hydrogen bond formed between His29 (purple) and Ser31 (yellow) from adjacent regulatory domains (blue) for the closed conformation of ^{V65P}*Tma*DAH7PS.

3.4 Targeted Molecular Dynamics

Targeted molecular dynamics (TMD) is a computational simulation where the atoms are guided towards a final "target" structure, by continuously decreasing the distance to the target via constraints.[101] For each time step the distance between the current coordinates and the target structure is calculated, following alignment of the two structures. The force applied to each atom is calculated from the potential energy gradient:

$$U_{total} = \frac{1}{2}\, k/N\, [RMS(t) - RMS^*(t)]^2$$

where the spring constant (k) is scaled down by the number of atoms (N) being simulated, $RMS(t)$ is the RMS distance between the current and target coordinates over time (t), and $RMS^*(t)$ evolves linearly from the initial RMSD to the final RMSD, reached at the final time step.

For TMD simulations the transition between conformational states is enforced independently of the height of the energy barriers that separate each conformer.[102] The constraints placed on the atoms of the simulated system, ensure the RMSD between the final structure adopted and the target structure, is virtually zero within the simulation run time.[103] This enables TMD to calculate a reaction pathway between two distinct conformational states. For a typical TMD transition, large-scale changes occur earlier in the simulation, followed by small-scale alterations; leading to some bias in the order of conformational rearrangement.[103] Despite this limitation TMD has proven a useful method for evaluating conformational transitions and has been used to assess allostery.[103]

3.4.1 Targeted Molecular Dynamics Simulations for *Tma*DAH7PS

TMD was utilised to investigate the transition from the fully open to fully closed and fully closed to fully open conformation of *Tma*DAH7PS. *Tma*DAH7PS crystal structures capturing the open (PDB: 1RZM) and closed (PDB: 3PG9) conformation of the enzyme were prepared as above, for the classical MD simulations; and were used as either the initial starting, or final target structure, depending on the direction of the TMD simulation. TMD simulations were run by applying a constant force (600 kcal/mol/Å^2) to the alpha carbon, of residues 1 to 80, which encompass the linker and regulatory domain of *Tma*DAH7PS.

The TMD simulation for the transition from the open to closed conformation of ^{WT}*Tma*DAH7PS was setup to occur over approximately 10 ns of classical MD simulation time. Protein backbone RMSD values indicate large deviations initially, corresponding to large global motions, which are characteristic of TMD. The RMSD fluctuates less as the closed conformation becomes more closely adopted, due to smaller local adjustments (Figure 3.13).

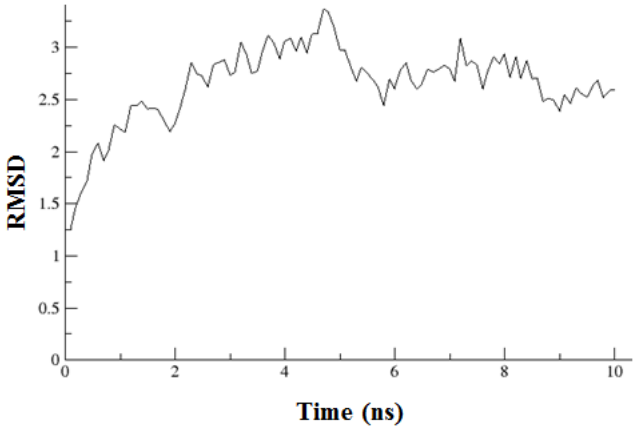

Figure 3.13: Backbone RMSD of ^{WT}*Tma*DAH7PS during TMD simulation for the transition from the open to closed enzyme conformation.

The TMD simulation for the transition from the closed to open conformation of ^{WT}*Tma*DAH7PS was setup to occur within the same time period as the open to closed TMD simulation. Protein backbone RMSD values indicate large deviations throughout the simulation (Figure 3.14). The decreased RMSD observed for the second half of the simulation corresponds to more pronounced local rearrangements for the open conformation.

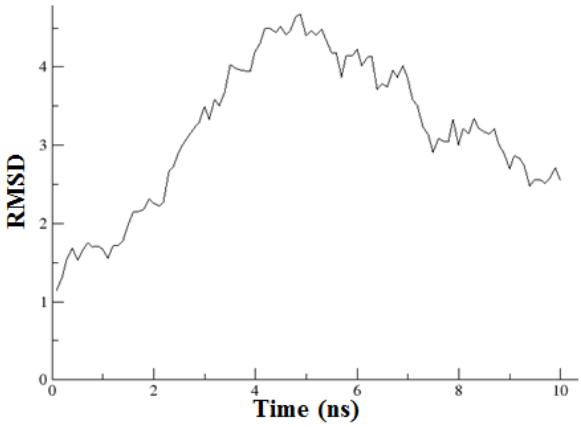

Figure 3.14: Backbone RMSD of ^{WT}*Tma*DAH7PS during TMD simulation for the transition from the closed to open enzyme conformation.

Visual examination of the TMD trajectories suggests a reasonably well defined path exists between the open and closed conformations of ^{WT}*Tma*DAH7PS (Figure 3.15). Some overlap of the conformational space explored during the classical MD and TMD simulations was observed, supporting the intermediate conformations characterised for ^{WT}*Tma*DAH7PS.

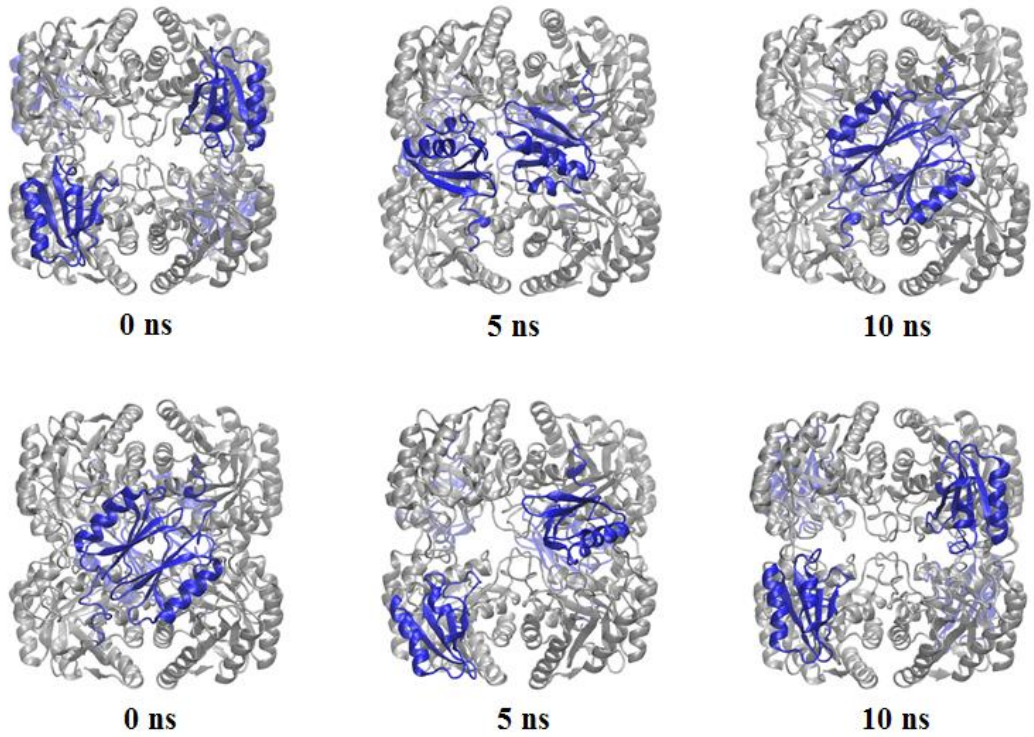

Figure 3.15: Conformational poses of ^{WT}*Tma*DAH7PS during a TMD simulation of the open to closed (above) and closed to open (below) conformation.

3.4.2 Dihedral Angles for Amino Acids of the Linker Region of *Tma*DAH7PS

For a protein, dihedral angles describe the rotation around the polypeptide backbone, about the C_α – N bond (φ) and the C_α – C bond (ψ). Over the course of the TMD simulations φ and ψ angles for the amino acids of the linker region were plotted, to explore which amino acids contribute to the movement of the regulatory domains, during the transition from the open to closed conformation of

^{WT}*Tma*DAH7PS. The φ and ψ values for each amino acid were averaged over the four protein chains of the tetramer and plotted for analysis.

The dihedral angle plots of the linker residues Arg64, Val65, Leu66 and Lys67 displayed large fluctuations in φ and/or ψ, over the course of the TMD simulation. The alterations in φ and ψ angles suggest each residue is potentially involved in the movement of the regulatory domains, during the transition from the open to closed conformation of ^{WT}*Tma*DAH7PS.

The ψ angle of Arg64 fluctuates throughout the TMD simulation, with a large rotation of the residue observed approximately midway through the transition from the open to closed conformation of ^{WT}*Tma*DAH7PS (Figure 3.16). The rotation of Arg64 is similar to Val65 and occurs within the same time period during the TMD simulation (Figure 3.17). This may indicate that Arg64 and Val65 are required to facilitate adoption of the closed conformation of the enzyme. Val65 undergoes the greatest fluctuation in φ and ψ angles compared to other linker residues. The dramatic changes in dihedral angles implicate Val65 in protein dynamics, which are likely essential in the transition from the open to closed conformation of ^{WT}*Tma*DAH7PS.

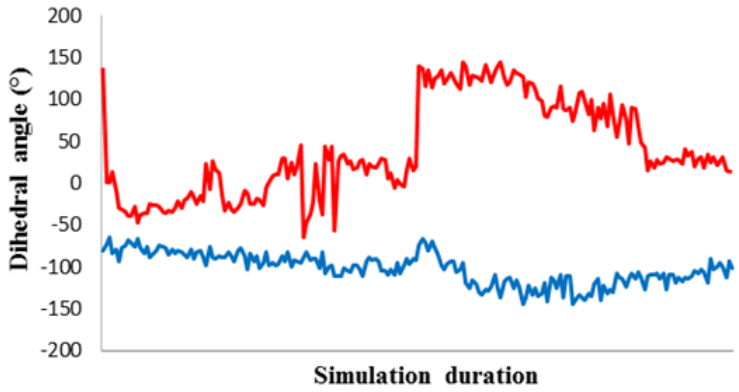

Figure 3.16: Averaged φ (blue) and ψ (red) angles for Arg64, plotted during the TMD simulation of ^{WT}*Tma*DAH7PS.

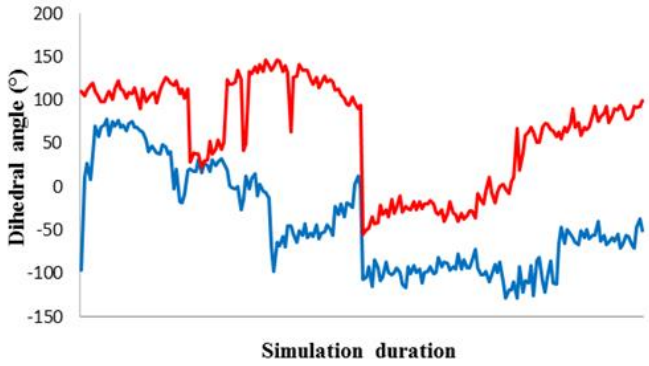

Figure 3.17: Averaged φ (blue) and ψ (red) angles for Val65, plotted during the TMD simulation of ^{WT}*Tma*DAH7PS.

Leu66 and Lys67 also exhibited fluctuations in φ and ψ angles. However, unlike Arg64 and Val65 there was no defined rotation observed midway through the TMD simulation. Leu66 undergoes dramatic changes in ψ angle as the closed conformation becomes more closely adopted (Figure 3.18). Interestingly Lys67 displayed sharp alterations in both φ and ψ angles around the midway point of the TMD simulation but otherwise these angles remained relatively constant (Figure 3.19). It seems likely that the minor fluctuations of Lys67 may be linked to the more dramatic changes occurring in the φ and ψ angles of Arg64, Val65 and Leu66.

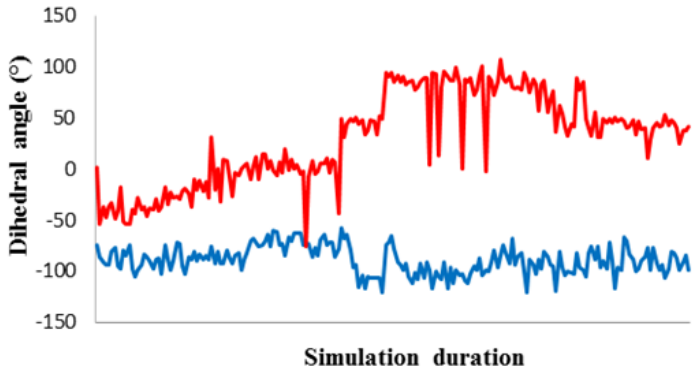

Figure 3.18: Averaged φ (blue) and ψ (red) angles for Leu66, plotted during the TMD simulation of ^{WT}*Tma*DAH7PS.

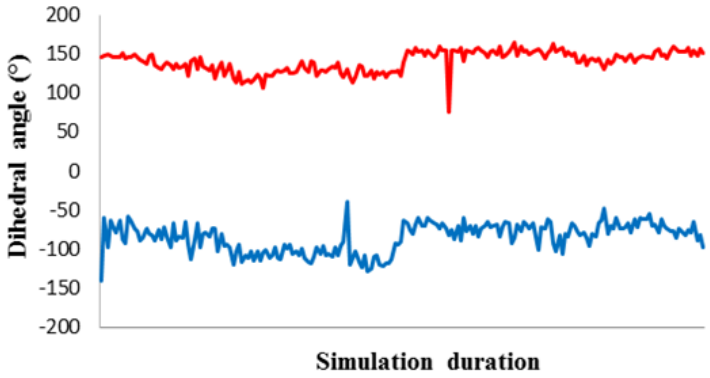

Figure 3.19: Averaged φ (blue) and ψ (red) angles for Lys67, plotted during the TMD simulation of ^{WT}*Tma*DAH7PS.

The φ and ψ angles of Arg64, Val65, Leu66 and Lys67 indicate each amino acid is important for linker dynamics. The movement of the regulatory domains is likely facilitated by each of the minor adjustments each amino acid undergoes, enabling the transition from the open to closed conformation of ^{WT}*Tma*DAH7PS.

Observations made from the MD and TMD simulations were used in combination with initial biochemical characterisation of ^{V65P}*Tma*DAH7PS to select additional linker variants to investigate. The mutations R64P, L66P and K67P were expected to act in a similar manner to the V65P linker substitution, where a reduction in flexibility of the linker region might lead to enhanced enzyme activity and/or a reduced response to Tyr. Based on the similarity of the dihedral angle plots of Arg64 and Val65, ^{R64P}*Tma*DAH7PS was selected over other possible linker residues for mutation.

3.5 Discussion

MD simulations are useful in the study of gross conformational changes and conformational flexibility, as biomolecules are able to be assessed on an atomic level. ^{WT}*Tma*DAH7PS has been studied previously via MD by Lang,[100] whose results corroborate those reported here for ^{WT}*Tma*DAH7PS, as well as providing context for the ^{V65P}*Tma*DAH7PS simulations. Classical MD simulations for the

open conformation of ^{WT}*Tma*DAH7PS, present greater deviations in protein backbone RMSD, relative to the open crystal structure, when compared to ^{V65P}*Tma*DAH7PS, indicative of greater conformational adjustment for ^{WT}*Tma*DAH7PS. Inspection of the simulation trajectories revealed conformational rearrangement of the regulatory domains, facilitated by the flexible inter-domain linker for ^{WT}*Tma*DAH7PS, with indication of the transition towards the closed conformation. Sampling of the fully closed conformation is likely unachievable via classical MD, due to the timescales able to be accessed. However, interaction of the ^{WT}*Tma*DAH7PS regulatory domains was observed within the simulation time achieved for ^{V65P}*Tma*DAH7PS, suggesting the restricted linker flexibility of ^{V65P}*Tma*DAH7PS, has greatly impacted the enzyme's ability to adopt the closed conformation, as no interaction of the regulatory domains was observed for ^{V65P}*Tma*DAH7PS.

The classical MD simulations for ^{V65P}*Tma*DAH7PS do not reveal adoption of the closed conformation or any conformation that resembles those adopted by ^{WT}*Tma*DAH7PS. MD simulations for the open conformation of ^{V65P}*Tma*DAH7PS instead reveal open conformations, where the regulatory domains proceed only so far towards the central axis of the tetramer, before swinging back towards more open conformations. Again, the time scale able to be accessed with classical MD, limits the conformations able to be observed. However, the conformational space explored by ^{V65P}*Tma*DAH7PS appears to be restricted by the less flexible linker region. The ^{V65P}*Tma*DAH7PS linker seems to facilitate the adoption of conformers, which do not easily access the closed enzyme conformation.

As MD simulations are limited by the time scale accessible, TMD simulations were utilised to accelerate conformational change, enabling computational observation within a practical time frame. The TMD simulation for the transition from the open to closed conformation of ^{WT}*Tma*DAH7PS, resulted in large deviations in protein backbone RMSD values initially, with less fluctuation as the closed conformation became more closely adopted. In comparison, the transition from the closed to open conformation of *Tma*DAH7PS, presented greater

deviations in protein backbone RMSD near the end of the simulation, as the open conformation became more closely adopted. The difference in protein backbone RMSD values for each TMD simulation highlights the inconsistences that occur for TMD simulations, due to propensity to perform large scale, global motions prior to small scale, local motions. The TMD simulations therefore represent only a possible route to the transition between the open and closed enzyme forms and do not represent the true transitional pathway, which connects the open and closed conformations.

Analysis of the dihedral angles for each amino acid of the linker region, implicates residues that impart flexibility to the linker. The flexibility of the linker is likely essential for linker dynamics and propagation of the allosteric signal between the regulatory domains and catalytic TIM barrel. In particular Val65 undergoes dramatic changes in φ and ψ angle over the course of the TMD simulation, so appears important for linker dynamics. The disruption of linker dynamics is therefore predicted to be the cause of the altered conformational sampling observed via classical MD, for ^{V65P}*Tma*DAH7PS. The altered conformers adopted by ^{V65P}*Tma*DAH7PS support the role of linker dynamics in the mechanism of allostery for *Tma*DAH7PS.

Chapter Four

Discussion of Linker Dynamics in Allosteric Regulation of *Thermotoga maritima* DAH7PS

4.1 Introduction

Perspective surrounding the role of inter-domain linkers has shifted from simple covalent attachment between discrete protein domains, to include the possibility of significant function, within complex allosteric signalling networks. Multiple studies support the role of linkers in allostery, reporting transitions between functional protein states that arise from linker dynamics.[104,105,106] Well studied examples include protein kinase A (PKA), which functions in cell signalling via phosphorylation of numerous cell targets, and the Hsp70 family of heat shock proteins (HSPs).

For PKA, the regulatory domain involved in ligand binding is connected by a flexible linker, which in the absence of cAMP is positioned over the active site of the catalytic subunit.[105] Upon cAMP binding the linker becomes more flexible and repositions, stabilising a conformation of the regulatory domain that binds weakly to the catalytic subunit. Stabilisation releases inhibition of the catalytic subunit, enabling kinase release .[105] The PKA linker, functions in population shift; altering the equilibrium that exists between the active and inactive enzyme forms, enabling a less populated state to be accessed by the conformational ensemble upon ligand binding.[107]

Hsp70 chaperone proteins also possess a dynamic linker, which functions in allosteric communication between the C-terminal substrate binding domain (SBD) and the N-terminal nucleotide binding domain (NBD).[106] The inter-domain linker

has been reported to act as a molecular switch, where upon substrate binding the linker is able to interact with the NBD, resulting in enhanced ATPase activity.[106] The Hsp70 linker, functions in allosteric coupling between the SBD and NBD; transducing allosteric signals between the domains in both directions, to control enzyme activity.

This chapter explores the role of linker dynamics in allosteric regulation of *Tma*DAH7PS, through discussion of the results presented in chapters two and three. When considered together, the results provide a more comprehensive understanding of the dynamic motions of *Tma*DAH7PS and the impact of linker dynamics on *Tma*DAH7PS allostery.

4.2 Role of Linker Dynamics in Allosteric Regulation of *Tma*DAH7PS

For *Tma*DAH7PS, the flexible inter-domain linker plays an important role in allosteric regulation of the enzyme. The inherent properties of the linker facilitate the adoption of two discrete conformers, which are significant for biological function. The transition between the open and closed conformation of *Tma*DAH7PS is possible due to the flexibility of the linker. However, re-distribution of the protein ensemble is controlled by dynamic properties, which are associated with ligand binding and transfer of the allosteric signal through the enzyme.

The linker variants ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS differ chemically from ^{WT}*Tma*DAH7PS, by single amino acid substitutions within the linker region. Individual substitutions alter linker flexibility and consequently dynamic properties of the variant enzyme. Characterisation of each *Tma*DAH7PS linker variant provides a means to assess the role of the linker, in *Tma*DAH7PS function and regulation.

*Tma*DAH7PS catalytic activity can be altered, via either decreased or increased linker flexibility, as evidenced by the greater k_{cat} values reported for ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS. k_{cat} is a measure of the number of substrate molecules converted to product, at each enzyme site per unit time. As ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS involve single amino acid substitutions within the linker, the enhanced k_{cat} may be attributed to a greater rate of substrate binding and/or product release, where the enhanced rate could arise from reduced sampling of the closed conformation of *Tma*DAH7PS.

The apparent reduction in sampling of the closed conformation in the absence of Tyr, for ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS supports the idea that for ^{WT}*Tma*DAH7PS, the closed conformation is accessible but is far less frequented than the open conformation when Tyr is absent. Alterations to linker flexibility appear to exacerbate the equilibrium bias, reducing sampling of the closed conformation, so that the open conformation is sampled more often and ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS are more catalytically active.

Decreased sampling of the closed conformation for ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS is supported by the classical MD simulations. MD simulations for ^{V65P}*Tma*DAH7PS do not predict adoption of the closed conformation, or any conformation that resembles the closed form. The altered flexibility of the linker appears to modify the conformational space explored by the enzyme, reducing the time spent sampling conformations that approximate the closed form.

The altered conformers adopted by ^{V65P}*Tma*DAH7PS, supports the role of linker dynamics in the mechanism of allostery for *Tma*DAH7PS. During the TMD simulations of the wild-type protein, Val65 underwent dramatic changes in φ and ψ angle, which were predicted to be related to linker dynamics associated with *Tma*DAH7PS allostery. The disruption to linker dynamics is likely the basis for altered conformational sampling, observed with classical MD, for ^{V65P}*Tma*DAH7PS.

The reduced sensitivity of ^{V65P}*Tma*DAH7PS to Tyr, also suggests the reduction in flexibility has disrupted linker dynamics, as the ability of the enzyme to adopt the closed conformation in the presence of Tyr has been impacted. The scattering data for ^{V65P}*Tma*DAH7PS, confirmed minor conformational change in the presence of Tyr, with overlay of the SAXS data producing similar SAXS profiles. The conformation adopted by ^{V65P}*Tma*DAH7PS, does not seem to enable full inhibition by Tyr. This suggests ^{V65P}*Tma*DAH7PS can either not access the closed conformation, or that the conformation accessed upon Tyr binding, is not the same as for ^{WT}*Tma*DAH7PS.

In contrast to ^{V65P}*Tma*DAH7PS, ^{P68A}*Tma*DAH7PS has improved sensitivity to Tyr. Like ^{WT}*Tma*DAH7PS, the allosteric regulation mechanism for ^{P68A}*Tma*DAH7PS seems to involve gross conformational change in the presence of Tyr. Overlay of the SAXS data in the absence and presence of Tyr is consistent with a change in conformation for ^{P68A}*Tma*DAH7PS. Fitting to the known structures of the wild-type enzyme using Crysol, suggests ^{P68A}*Tma*DAH7PS adopts virtually identical conformations to the open and closed enzyme structures in response to Tyr binding.

Increased flexibility of the ^{P68A}*Tma*DAH7PS linker has improved enzyme activity and the potency of the allosteric response. The increased flexibility seems to enable adoption of both the open and closed enzyme conformations. For ^{P68A}*Tma*DAH7PS, linker dynamics result in more exclusive sampling of both the open and closed enzyme conformations, dependent on the absence or presence of Tyr. This suggests the equilibrium between the two discrete conformers remains intact but is strongly biased relative to Tyr, for ^{P68A}*Tma*DAH7PS.

4.3 Conclusion

The flexibility of the linker is integral to the mechanism of allostery employed by *Tma*DAH7PS. Alteration to linker flexibility can significantly impact linker

dynamics, which are essential for the transition between discrete functional states. The dynamic properties of the linker control the position of the equilibrium between the open and closed enzyme forms. Adjusting linker flexibility impacts the equilibrium, generating variant enzymes with distinct catalytic and regulatory properties, which arise from the individual dynamic population ensembles.

Chapter Five

Summary of Thesis and Future Work

Protein molecules are highly dynamic species, which exist as an ensemble of protein conformations that readily interconvert, via dynamic fluctuations. The role of inter-domain linkers in allostery has become more recognised over the past decade, with multiple studies reporting transitions between functional protein states, attributed to linker dynamics.[104-107] In this thesis the biochemical characterisation of *Tma*DAH7PS and several *Tma*DAH7PS linker variants is described. The results presented here for ^{WT}*Tma*DAH7PS align well with other studies performed on this enzyme, both within the Parker research group and by other research groups; providing context for each of the *Tma*DAH7PS linker variants discussed during this work.[33,82] The *Tma*DAH7PS linker variants ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS confirm the importance of linker flexibility for adoption of discrete functional conformers, and the role of linker dynamics in conformational selection and allosteric propagation.

5.1 Modifying Linker Flexibility Impacts *Tma*DAH7PS Activity and Regulation by Tyrosine

The linker variants ^{V65P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS have increased k_{cat} in comparison to ^{WT}*Tma*DAH7PS, indicating that either a decrease or increase in linker flexibility can impact *Tma*DAH7PS activity. The enhanced catalytic rate is predicted to be caused by a shift in the equilibrium position between discrete functional states. The equilibrium may favour the open conformer, resulting in decreased sampling of the closed conformation and therefore an increase in enzyme activity. Decreased sampling of the closed conformation is supported by

the MD simulations, which confirm the adoption of open-like conformers by ^{V65P}*Tma*DAH7PS. Modifying linker flexibility appears to adjust the conformational ensemble, limiting the conformational space explored by the enzyme.

^{R64P}*Tma*DAH7PS and ^{V65P}*Tma*DAH7PS are less sensitive to inhibition by Tyr, suggesting decreased linker flexibility has disrupted linker dynamics, impairing the ability of each enzyme to adopt the closed conformation. The response of ^{R64P}*Tma*DAH7PS to Tyr was more significant than ^{V65P}*Tma*DAH7PS however, accentuating the importance of Val65 in the dynamic motions associated with adoption of the open and closed conformations. In contrast ^{P68A}*Tma*DAH7PS was more sensitive to inhibition by Tyr, with increased flexibility of the linker aiding adoption of the closed conformation when Tyr was present. The experimental results suggest the dynamic motions required to transition between the open and closed conformations, remain intact for ^{P68A}*Tma*DAH7PS.

5.2 Linker Dynamics Facilitate Allosteric Regulation of *Tma*DAH7PS

The allosteric regulation mechanism of *Tma*DAH7PS is significantly impaired by reducing the flexibility of the linker. The absence of conformational change in the presence of Tyr for ^{V65P}*Tma*DAH7PS was confirmed by SAXS, where overlay of the SAXS data collected in the absence and presence of Tyr, generated similar SAXS profiles, which deviate little from one another. The conformation adopted by ^{V65P}*Tma*DAH7PS does not enable full inhibition by Tyr. It appears that linker dynamics have been disrupted for ^{V65P}*Tma*DAH7PS, so that the closed conformation adopted by ^{WT}*Tma*DAH7PS upon Tyr binding, is unable to be accessed by ^{V65P}*Tma*DAH7PS.

For ^{P68A}*Tma*DAH7PS the allosteric regulation mechanism involves gross conformational change upon Tyr binding. The increased flexibility of the linker

enables adoption of both the open and closed enzyme conformations. Fitting to the existing structures of the wild-type protein by Crysol, matches the SAXS profiles for ^{P68A}*Tma*DAH7PS in the absence and presence of Tyr, to the open and closed crystal structures with high fidelity. ^{P68A}*Tma*DAH7PS linker dynamics facilitate the adoption of both the open and closed conformations. However, the increased flexibility has influenced dynamic motions, so that discrete functional conformations are sampled more exclusively, dependent on the absence or presence of Tyr.

5.3 Future Work

The role of linker dynamics in allosteric regulation of *Tma*DAH7PS could be further elucidated by several key studies, some of which were attempted but were unable to be achieved within the timeframe of this thesis. The studies outlined here may provide key evidence, required to answer the following questions.

What is the Structure of ^{V65P}*Tma*DAH7PS?

The structure of ^{V65P}*Tma*DAH7PS in the presence of Tyr was sought for comparison to the closed crystal structure of ^{WT}*Tma*DAH7PS. For ^{WT}*Tma*DAH7PS, the crystal for the closed structure was obtained by sitting-drop vapour diffusion, grown from a reservoir containing 20% (w/v) polyethylene glycol 3350, 0.2 M ammonium nitrate and 0.02% (w/v) sodium azide.[82] Attempts were made to crystallise ^{V65P}*Tma*DAH7PS in the absence and presence of Tyr, using the hanging-drop vapour diffusion method, screening similar conditions to ^{WT}*Tma*DAH7PS. Unfortunately, the initial screens proved unsuccessful but higher throughput screening may enable detection of conditions under which the protein will crystallise. Structural information for ^{V65P}*Tma*DAH7PS would improve the validity of the results presented here, particularly Crysol fitting of the ^{V65P}*Tma*DAH7PS SAXS data to the open and closed crystal structures and the

MD simulations, which would benefit from the ^{V65P}*Tma*DAH7PS crystal structure as opposed to the ^{WT}*Tma*DAH7PS crystal structure.

Does ^{V65P}*Tma*DAH7PS Bind Tyrosine?

Isothermal titration calorimetry (ITC) is a technique that enables determination of the thermodynamic parameters of interactions in solution.[108] As ITC provides a direct measure of the heat changes associated with ligand binding in solution, the technique was employed to study binding interactions between Tyr and *Tma*DAH7PS and the *Tma*DAH7PS linker variants; with particular interest in the binding properties of ^{V65P}*Tma*DAH7PS. Although unsuccessful, the ITC analysis attempted during this research is presented here, to provide a basis for future optimisation of the *Tma*DAH7PS – Tyr system, which will be invaluable for determination of the molecular detail surrounding *Tma*DAH7PS allostery.

For ^{WT}*Tma*DAH7PS a range of protein and ligand concentrations were trialled with limited success. Best results were obtained when a 1 mM solution of Tyr was titrated into 155 µM ^{WT}*Tma*DAH7PS and fit to a one-site binding model. However, low sigmoidicity of the binding curve indicated the fit was inadequate to generate usable binding data. Sigmoidicity of the binding curve may be improved by optimising the relative concentrations of protein and ligand. From analysis of the conditions attempted previously, a lower concentration of Tyr or smaller injection volume may increase the number of peaks observed before saturation, improving sigmoidicity of the binding curve and accuracy of the thermodynamic parameters.

A number of protein and ligand concentrations were also trialled for ^{V65P}*Tma*DAH7PS, with little improvement. A more extensive range of both protein and ligand concentrations were evaluated for ^{V65P}*Tma*DAH7PS. However, the binding curves generated had little sigmoidicity, when fit to a one-site binding model. Tyr concentrations from 0.3 to 3 mM were assessed, at a variety of protein concentrations. The most promising condition comprised 0.5 mM Tyr and 214

μM ^{V65P}*Tma*DAH7PS; although it is possible a higher Tyr or lower protein concentration may improve the quality of fit to a one-site binding model.

Initial characterisation weakly suggests Tyr binds to ^{V65P}*Tma*DAH7PS, with a K_d within the same concentration range as ^{WT}*Tma*DAH7PS obtained. However, with the quality of data at present, an accurate K_d is unable to be reported. It is unlikely that optimisation of the protein and ligand concentrations alone, will improve ^{WT}*Tma*DAH7PS or ^{V65P}*Tma*DAH7PS ITC data. The *Tma*DAH7PS – Tyr system proves difficult to assess by ITC and this may be for a number of reasons, including mutually exclusive conditions required to optimise K_d and ΔH. Factors to consider when moving forward include the heat of dilution, buffer selected, pH and temperature of the experiment.

A relatively large heat of dilution was observed that in part, is likely due to mixing effects associated with the properties of Tyr but also factors, such as the buffer Tyr is dissolved in. Buffer mismatch between the macromolecule in the sample cell and the ligand injected, can result in large heats being produced, which mask potential binding interactions. Extensive dialysis of *Tma*DAH7PS enzymes into the appropriate buffer was performed. However, purification of each enzyme, using the selected buffer for ITC, may be more effective in preventing buffer mismatch between *Tma*DAH7PS and Tyr.

Further optimisation of the buffer and pH employed may also improve quality of the ITC data. Buffers with a low enthalpy of ionisation should be utilised to minimise artificial heat production.[108] The temperature at which the experiment is conducted can also impact the heats produced, so the experiment was repeated at 60 °C. Unfortunately, no improvement was observed. As *Tma*DAH7PS is a hyper-thermophile elevated temperatures may enhance quality of the binding data obtained. However, additional evaluation at a range of temperatures is required.

Elucidation of the role of the inter-domain linker, in the allosteric regulation mechanism of *Tma*DAH7PS would benefit from molecular detail regarding the binding interactions between Tyr and *Tma*DAH7PS. In particular the contribution

of linker dynamics to *Tma*DAH7PS allostery would be supported by ligand binding information, which would provide more definitive evidence for conformational selection and allosteric propagation, resulting from perturbation of linker dynamics.

Can the Open/Closed Response be Triggered Without the TIM Barrel Scaffold?

The creation of a synthetic dimer between two of the regulatory domains of *Tma*DAH7PS, could confirm whether the open/closed response is possible without the TIM barrel scaffold. If the regulatory domains are connected together with a flexible inter-domain linker, the synthetic dimer may act as a functional unit, which in the absence and presence of Tyr, could switch between the open and closed forms. Linkers of varying lengths could be trialled to determine whether the open/closed response could be triggered for the regulatory domains. Synthesis of an open/closed functional unit may enable transfer to a completely new protein scaffold, imparting allosteric regulation to the protein.

Chapter Six

Experimental Methods

6.1 General Methods

6.1.1 Protein Structures

The PyMOL Molecular Graphics System (Version 1.8, Schrödinger LLC) was used to generate all protein structure figures.[109]

6.1.2 Sequence Alignments

Multiple amino acid sequence alignments were generated using the Clustal Omega web server (http://www.ebi.ac.uk/Tools/msa/clustalo/).[110]

6.1.3 Purified Water

All water used for experimental procedures was purified by a Millipore Milli-Q water system and is referred to as Milli-Q. Milli-Q required for molecular biology was autoclaved prior to use.

6.1.4 Chelexed Milli-Q

Chelex® 100 resin (Bio-Rad) was used to remove metal ions from solution. The manufacturer's instructions for the batch method were followed, where the resin was added to the Milli-Q solution, left to stir overnight and filtered (0.2 μm) prior to use.

6.1.5 pH Determination

The pH of all buffer solutions was measured using a Denver UB-10 Ultrabasic pH meter at room temperature. Desired pH was achieved via the addition of either HCl or NaOH, at a concentration of 1 M or 10 M, to the buffer solution.

6.1.6 Antibiotic Stocks

Ampicillin (100 mg/mL) stock solutions were prepared in Milli-Q, filtered (0.2 μm), aliquoted into micro-centrifuge tubes under sterile conditions and stored at -80 °C.

6.1.7 Media

Super optimal broth with catabolite repression (SOC) media (2% (w/v) tryptone, 0.05% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose) was prepared in Milli-Q, with the addition of each component, excluding $MgCl_2$ and glucose. The mixture was autoclaved at 121 °C, 15 psi for a minimum of 20 minutes and aliquoted into sterile micro-centrifuge tubes before the addition of filter sterilised $MgCl_2$ and glucose immediately prior to use.

Cell cultures were grown in LB media (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) prepared in Milli-Q at a concentration of 20 g/L. LB media was

sterilised by autoclaving at 121 °C, 15 psi for a minimum of 20 minutes prior to use. LB agar plates were prepared via the addition of agar (25 g/L) to LB. LB agar was heated to boiling, followed by incubation at 60 °C for one 1 hour. Ampicillin (100 mg/mL) was added immediately prior to pouring.

6.1.8 Chemically Competent Cells

All chemically competent cells were commercially prepared and supplied (Invitrogen or Novagen).

6.1.9 Transformation

Chemically competent cells (50 µL aliquots) were thawed on ice prior to the addition of purified plasmid (2 µL). The cells were incubated on ice for 30 minutes before a 30-45 second heat shock at 42 °C. The cells were incubated on ice for a further 2 minutes, prior to the addition of SOC media (250 µL) and incubated for 1 hour at 37 °C, 180 rpm. Transformed cells (50-200 µL) were spread on a pre-warmed LB agar plate, supplemented with ampicillin and incubated overnight at 37 °C.

6.1.10 Polymerase Chain Reaction

Polymerase chain reactions (PCRs) were performed by following the standard protocol for Phusion® High Fidelity DNA Polymerase (Thermo Scientific). The PCR reaction mixture (50 µL) consisted of 1x reaction buffer, deoxyribonucleotide triphosphates (dNTPs) (200 µM), forward and reverse primers (0.5 µM), template DNA (50-250 ng/gDNA) and Phusion® High Fidelity DNA Polymerase (10 µL). A two or three step thermocycling protocol was employed for each reaction.

6.1.11 Agarose Gel Electrophoresis

Separation of DNA fragments was achieved via agarose gel electrophoresis. Agarose gels were prepared using a solution of 1% (w/v) LE agarose (Seakem) in 50x Tris base-acetic acid-EDTA (TAE) buffer (40 mM Tris-HCl, 20 mM glacial acetic acid, 2 mM EDTA). The solution was heated until the agarose was completely dissolved then incubated at 60 °C, before the addition of SYBR® Safe DNA Gel Stain (Invitrogen). Samples were mixed with 6x loading buffer (60 mM Tris-HCl, 60 mM EDTA, 0.02% (w/v) orange G, 0.05% (w/v) xylene cyanol FF, 60% glycerol) before loading to gel wells. A 1 kB Plus DNA ladder (Invitrogen) was run alongside the samples for size comparison. Hand-cast gels were run in 1x TAE buffer using a Mini-Sub cell GT (Bio-Rad), at 85 V for 45 minutes. Images were captured using a Molecular Imager® Gel Doc™ XR, visualised under UV (302 nm).

6.1.12 Restriction Enzyme Digest

High fidelity restriction enzymes (New England Biolabs) were utilised to digest circular vectors. Plasmid DNA (1 µg) was combined with 10x Cutsmart® Buffer, restriction enzyme (1 µL) and Milli-Q to a total volume of 50 µL before the restriction enzyme digest mixture was incubated at 37 °C for 1 hour. The restriction enzymes were then denatured at 80 °C for 20 minutes. The digested vector was purified by gel electrophoresis prior to use.

6.1.13 Plasmid Extraction and Purification

A 5 mL LB preculture, supplemented with ampicillin was inoculated with a single colony of either Stellar™ or *E. coli* BL21*(DE3) chemically competent cells (Clontech), which had been transformed with the desired plasmid. Cells were grown overnight at 37 °C, 180 rpm. The cells were harvested and plasmids

extracted using a Nucleospin® Plasmid Easy Pure Kit (Clontech) and the concentration of purified plasmid measured using a Nanodrop-1000 spectrophotometer under UV (260 nm).

6.1.14 DNA Sequencing

DNA sequencing services were provided by Macrogen. For the sequencing reactions approximately 200 ng of plasmid and 3.2 µM sequencing primers were supplied.

6.1.15 Glycerol Stocks

Glycerol stocks were prepared by combining an aliquot of preculture growth with 20% (v/v) glycerol, under sterile conditions. The glycerol stock was mixed thoroughly before transfer to a sterile micro-centrifuge tube, which was snap frozen in liquid nitrogen and stored at -80 °C.

6.1.16 *E. coli* Cell Culture Growth

For *E. coli* cell cultures a preculture was prepared from 50 mL of LB media, supplemented with ampicillin and inoculated with a glycerol stock scraping. The preculture was incubated overnight at 37 °C, 180 rpm. The preculture was used to inoculate 1 L of LB media, supplemented with ampicillin. 1 L cultures were grown in baffled flasks at 37°C, 180 rpm until mid-logarithmic phase (OD_{600} 0.4-0.8 AU). Protein expression was induced with IPTG (Roche), which was added to cultures at a final concentration of 1 mM. Cells were harvested approximately 4 hours after incubation with IPTG (37 °C, 180 rpm) by centrifugation (12,000 *g*, 4 °C, for 20 minutes) and the cell pellets stored at -80 °C.

6.1.17 Cell Lysis

The cell pellet was re-suspended in approximately 20 mL of lysis buffer (10 mM 1,3-(tris(hydroxyl-methyl)-methylamino)propane (BTP), 1 mM EDTA, 100 mM KCl, 200 µM PEP, pH 7.3) and cells lysed by sonication, using an Omni-Ruptor 4000 Ultrasonic Homogenizer (Omni International). The cell pellets were sonicated in a beaker surrounded by ice and subjected to 3x 5 minute rounds at 80% power, 30% pulsation followed by 3x 5 minute rounds at 80% power, 40% pulsation. The cell debris was removed by centrifugation (40,000 *g*, 4 °C for 30 minutes) and the supernatant incubated with 2 µL Benzonase® Nuclease (Sigma Aldrich) for 10 minutes.

6.1.18 Heat Treatment

Heat treatment of the cell lysate was performed using a heat block at 70 °C, for 30 minutes. The cell lysate was aliquoted into test tubes and covered with Parafilm M® prior to heating. Following the heat treatment test tubes were left to cool to room temperature, before dilution with lysis buffer to improve protein collection. Heat debris was removed by centrifugation (40,000 *g*, 4 °C for 30 minutes) and the supernatant collected for further purification.

6.1.19 Fast Protein Liquid Chromatography

Fast protein liquid chromatography (FPLC) was performed using either a Bio-Rad Biologic Protein Chromatography system or an ÄKTApurifier™ (GE Healthcare), at 4 °C or room temperature to purify all proteins. All solvents and buffers were filtered using a 0.2 µm filter (Millipore) prior to use. Protein samples were also filtered (0.2 µm) before loading to a 10 mL or 50 mL Superloop™ (GE Healthcare). The eluant was collected in 2 mL fractions within 96-well plates.

6.1.20 Hydrophobic Interaction Chromatography

For HIC the protein sample was slowly combined with $(NH_4)_2SO_4$ (1 M), immediately prior to being loaded onto a Source™ 15Phe® (GE Healthcare) column, pre-equilibrated with high salt buffer (10 mM BTP, 1 mM EDTA, 1 M $(NH_4)_2SO_4$, pH 7.3). The HIC procedure was performed at room temperature following a step-wise method, where the protein was eluted from the column with low salt buffer (10 mM BTP, 1 mM EDTA, pH 7.3), as the concentration of $(NH_4)_2SO_4$ was lowered from 1 M to 0 M. Fractions exhibiting UV absorbance at 280 nm were analysed by SDS-PAGE and those containing the protein of interest pooled and concentrated, using a 20 mL 10 kDa molecular weight cut-off (MWCO) device (Sartorius Stedim Biotech).

6.1.21 Size-Exclusion Chromatography

SEC was performed using either a Superdex S200 26/60 (GE Healthcare) or Superdex S200 HR 10/300 column (GE Healthcare) pre-equilibrated with buffer (10 mM BTP, 100 mM KCl, 200 µM PEP, 1 mM EDTA, pH 7.3). SEC was conducted at 4 °C and protein eluted from the column dependent on biomolecule size. Fractions exhibiting UV absorbance at 280 nm were analysed by SDS-PAGE and those containing the protein of interest pooled and concentrated, using a 20 mL 10 kDa MWCO device.

6.1.22 Polyacrylamide Gel Electrophoresis

SDS-PAGE was performed using a Bolt® 10% Bis-Tris Plus precast protein gel (Invitrogen) run in a Mini Gel Tank (Invitrogen). Samples were boiled prior to loading of gel wells and a Novex® Sharp pre-stained protein standard (Invitrogen)

run as a molecular weight marker. Gels were run in MOPS buffer at 200 V, for 32 minutes. For protein visualisation gels were stained with Coomassie Brilliant Blue R250 in 50% (v/v) methanol, 10% (v/v) acetic acid for approximately 30 minutes, followed by de-staining in 40% (v/v) methanol and 10% (v/v) acetic acid for 30 minutes. Images were captured using a Molecular Imager® Gel Doc™ XR, visualised under UV (302 nm).

6.1.23 Protein Concentration and Buffer Exchange

For protein concentration during and after purification, a 20 mL 10 kDa MWCO device was utilised, which had been washed with 2-3 volumes of Milli-Q. Proteins were buffer exchanged via dilution with the desired buffer and repeatedly concentrated by centrifugation (4000 *g*, 4 °C).

6.1.24 Protein Concentration Determination

Protein concentration was determined by measuring the UV absorbance at 280 nm for a 2 µL sample, with a Nanodrop ND-1000 spectrophotometer (Thermo Scientific). The protein extinction coefficient was calculated using the ExPASY ProtParam tool (http://web.expasy.org/protparam/).

6.1.25 Protein Storage

Purified protein was separated into 50 µL, 100µL or 200 µL aliquots; snap frozen in liquid nitrogen and stored at -80 °C.

6.1.26 Molecular Weight Determination

Protein molecular weight was measured by electrospray ionisation mass spectrometry using a Bruker maXis^{TM} 3G (Bruker Daltonics). Protein samples were buffer exchanged into Milli-Q and diluted to a concentration of 1 mg/mL for analysis.

6.1.27 Enzyme Kinetics Assay

DAH7PS activity was assessed using a continuous spectrophotometric assay, performed at 60 °C. The consumption of PEP was monitored at 232 nm (ε = 2.6x 10^3 M^{-1} cm^{-1}) using a Varian Cary 100 UV Visible Spectrophotometer. Measurements were made in triplicate using a quartz cuvette, with a 1 cm path length. Standard reaction mixtures contained PEP (200 µM), E4P (300 µM) and $MnSO_4$ (100 µM) in BTP buffer (50 mM BTP, pH 7.3), with enzyme (2 mg/mL) to a total volume of 1 mL. BTP buffer was made up in Chelexed Milli-Q, which was then used to prepare PEP and E4P solutions. $MnSO_4$ and Tyr solutions were made up in Chelexed Milli-Q. The reaction mixture comprising PEP, $MnSO_4$ and BTP buffer was incubated for 7 minutes at 60 °C prior to the addition of enzyme. The reaction mixture was then incubated for an additional 2 minutes before the assay was initiated by the addition of E4P. Initial rates of reaction were determined by least squares fitting to the initial rate data. *K*m and k_{cat} values were determined by fitting the initial rates to the Michaelis-Menten equation using Grafit software (Erithacus).

6.1.28 Substrate Concentration Determination

A standard assay was performed in triplicate to determine the concentration of PEP and E4P stock solutions. To assess the concentration of either substrate, a reaction mixture containing one substrate in excess, with the other limiting,

$MnSO_4$ (100 µM) and enzyme in BTP buffer; was initiated by the addition of E4P. Upon completion of the reaction the change in absorbance was measured and the Beer-Lambert Law applied to determine the concentration of either PEP or E4P stock solutions (ε = 2.8x 10^3 M^{-1} cm^{-1} at 25 °C or ε = 2.6x 10^3 M^{-1} cm^{-1}at 60 °C, *l* = 1 cm).

6.1.29 Feedback Inhibition Studies

Tyr and Phe stock solutions (20 mM) were made up in Chelexed Milli-Q and added to standard assay reaction mixtures to give a total reaction volume of 1 mL.

6.2 Chapter Two Methods

6.2.1 PCR Amplification of ^{WT}*Tma*DAH7PS

Before commencing this research, the gene for *Tma*DAH7PS (locus tag TM0343) had been amplified from purified *T. maritima* MSB8 (DSM 3109) genomic DNA, using primers designed to introduce NdeI and BglII recognition sites. The 1036 bp PCR product was ligated into the pT7-7 vector, digested with NdeI and BamHI.[83] The pT7-7 vector was transferred into *E. coli* XL1-Blue cells and the resulting construct transformed into *E. coli* BL21*(DE3) cell lines for protein expression. Sequencing analysis confirmed the presence of the ^{WT}*Tma*DAH7PS sequence, within the expression plasmid.

6.2.2 Site-directed Mutagenesis

The linker variants ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS were constructed by site-directed mutagenesis using a Quik-Change® Lightning

Site-Directed Mutagenesis Kit (Stratagene). ^{V65P}*Tma*DAH7PS was constructed prior to the beginning of this research however the same protocol was followed to generate the R64P and P68A *Tma*DAH7PS linker mutations. The pT7-7 *Tma*DAH7PS plasmid was used as a double stranded template and the amino acid substitutions (R64P or P68A) introduced using synthetic oligonucleotide primers, harbouring the desired mutations; complementary to each strand of the pT7-7 *Tma*DAH7PS plasmid. The forward and reverse primers for the R64P and P68A *Tma*DAH7PS linker mutations are shown below, with codons representing the mutation shown in bold.

^{R64P}*Tma*DAH7PS

F: 5'-tgtaaggtttgagca**ccg**gcacaacgctttccacgc-3'

R: 5'-gcgtggaaagcgttgtgc**cgg**tgctcaaaccttaca-3'

^{P68A}*Tma*DAH7PS

F: 5'-gagaaacgagtttgtaa**gct**ttgagcaccctcacaac-3'

R: 5'-gttgtgagggtgctcaa**agc**ttacaaactcgtttctc-3'

The Quik-Change® Lightning Site-Directed Mutagenesis procedure involved extension of mutagenic primers by Phusion® High Fidelity DNA polymerase during temperature cycling, to generate a nicked mutant plasmid. The PCR product was treated with DpnI to digest parental DNA, before transformation into Stellar™ competent cells. Transformants were grown overnight at 37 °C and the plasmids extracted using a Nucleospin® Plasmid Easy Pure Kit for DNA sequencing. Upon sequence confirmation, constructs were transformed into *E. coli* BL21*(DE3) cell lines for protein expression.

6.2.3 Kinetic Characterisation

The standard enzyme kinetics assay was employed to determine the K_m^{PEP} and K_m^{E4P} for ^{WT}*Tma*DAH7PS, ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS. For determination of K_m^{PEP}, E4P concentration was fixed at 300 µM, while PEP concentration was varied from 20 µM to 200 µM. For determination of K_m^{E4P}, PEP concentration was fixed at 200 µM, while E4P concentration was varied from 20 µM to 300 µM.

6.2.4 Enzyme Activity in Presence of Inhibitors

The standard enzyme kinetics assay was employed to determine enzyme activity in the presence of Tyr or Phe for ^{WT}*Tma*DAH7PS, ^{V65P}*Tma*DAH7PS, ^{R64P}*Tma*DAH7PS and ^{P68A}*Tma*DAH7PS. PEP concentration was fixed at 200 µM and E4P concentration at 300 µM, while the concentration of Tyr or Phe was varied from 0 µM to 400 µM. Initial rates for each ligand concentration were compared to those of standard assay conditions to determine the loss of activity. Assays were performed in triplicate.

6.2.5 Differential Scanning Fluorimetry

DSF was employed to measure protein melt temperatures using an iCycler iQ5 Multicolour Real-Time PCR Detection System (Bio-Rad). Protein samples (1 mg/mL) with buffer (50 mM BTP, pH 7.3), 20x SYPRO® Orange dye and 1 mM Tyr (where necessary) were aliquoted as 25 µL triplicates, into a 96-well plate. A blank was prepared for each sample, by substituting protein for buffer. The melt proceeded in 0.2 °C increments from 20 °C to 95 °C, with a dwell time of 20 seconds after each temperature rise. Fluorescence was measured at the end of each dwell time.

6.2.6 Differential Scanning Fluorimetry Data Analysis

DSF data analysis was performed using an Excel spreadsheet with a custom VBA-scripted macro to toggle and dynamically display melt and derivative curves, for each protein sample. Melt temperatures were calculated as the temperature of maximum inflection of the melting curve, after subtracting absorbance of the sample blank. Error bars associated with DSF plots represent the standard deviation between triplicate measurements.

6.2.7 Small Angle X-ray Scattering Data Collection

SAXS data was collected at the Australian Synchrotron SAXS/WAXS beamline equipped with a Pilatus detector. The wavelength of the X-rays was 1.0332 Å and the sample detector distance set at 1600 mm. Scattering data was collected from enzymes (10 mg/mL) following elution from a SEC column (Superdex 200 5/150) pre-equilibrated with buffer (10 mM BTP, 100 mM KCl, 200 µM PEP, pH 7.3) with or without Tyr (1 mM). Two-dimensional intensity plots from the peak of the size-exclusion chromatography run were radially averaged, normalised to sample transmission and the background subtracted.

6.2.8 Small Angle X-ray Scattering Data Analysis

The scatterbrain software (Australian Synchrotron) was utilised to process all primary SAXS data collected. Guinier fits were plotted using PRIMUS.[85] Indirect Fourier transformation was performed using GNOM to generate the P(r) function, providing both the relative probability of distances between scattering centres and the maximum dimension of scattering (D_{max}). Theoretical scattering curves were generated from atomic coordinates (PDB 1RZM and 3PG9) using CRYSOL.[86] The quality of CRYSOL fitting was assessed by χ^2 values.

6.3 Chapter Three Methods

6.3.1 Open and Closed System Preparation

*Tma*DAH7PS crystal structures capturing both the open form co-crystallised with PEP and E4P (PDB: 1RZM) and the closed form co-crystallised with Tyr (PDB: 3PG9) were used as starting points for the MD simulations. Tetrameric structures were generated from the associated dimer in the asymmetric unit, based on crystallographic symmetry. Preparation of the open form involved the removal of PEP and E4P molecules to enable protein dynamics to be assessed in the absence of ligands. Replacement of active site Cd^{2+} ions with Mn^{2+} ions was performed using COOT[111] and parameters for Mn^{2+} added to the CHARMM all atom parameter set 22.[88] The protonation state for ionisable residues was calculated using MCCE[112] for histidine and PROPKA[113] for all other residues, with the results averaged per protein chain. The open system was solvated in a TIP3 water box of dimensions 109 Å X 97.2 Å X 98.9 Å, while the closed system had dimensions 99.1 Å X 102 Å X 101 Å. The net charge of the system was neutralised via the addition of Na^+ and Cl^- ions, which were added with a minimum 5 Å distance between individual ions and between the enzyme and each ion.

6.3.2 Molecular Dynamics Simulations

NAMD[88] was utilised to carry out all MD simulations, using the CHARMM all atom parameter set 22. MD simulations were run using 1024 cores of the IBM Blue Gene/P computer from the high performance supercomputing and services facility, at the University of Canterbury. Electrostatic interactions were calculated using the PME method, with a Van der Waals cut-off of 12 Å. Each MD simulation was performed at 1 atmosphere and 363 K, the known optimal temperature for *T. maritima* activity. For each simulation an energy minimisation

of 5000 steps, using the conjugate gradient energy minimisation method was run. Time steps of 2 fs were utilised and a trajectory written out every 100 ps.

6.3.3 Targeted Molecular Dynamics Simulations

TMD simulations were run using the open and closed systems as either the initial starting, or final target structure, depending on the direction of the simulation. TMD simulations were run for the open to closed form and closed to open form, by applying a force constant of 600 kcal/mol/Å^2 to the alpha carbon of residues 1 to 80 of *Tma*DAH7PS. TMD simulations were run over 5,000,000 steps, with the relevant open or closed crystal structure used as the final target, with a RMSD of 0.2 Å set for the final TMD step.

6.3.4 Trajectory Analysis

All simulations were prepared, displayed and analysed using the molecular visualisation program VMD.[114] For both the open and closed conformation of *Tma*DAH7PS Val65 was mutated to Pro (V65P) *in silico*, using mutator plug-in (version 1.3). Trajectories were aligned in VMD with respect to the TIM barrel of *Tma*DAH7PS and RMSD and dihedral angles calculated using Gromacs 4.5.[115]

6.4 Chapter Five Methods

6.4.1 Crystallisation Trials

Manual screening for crystallisation conditions, via hanging-drop vapour diffusion in 24 well plates, was performed using 1 µL drops of protein and 500 µL of appropriate well solution. ^{V65P}*Tma*DAH7PS concentrations tested included 4

mg/mL, 6 mg/mL and 10 mg/mL. Crystal trays were prepared at room temperature before storage at 20 °C. Crystal trays and pipette tips were flushed with nitrogen prior to use.

6.4.2 Isothermal Titration Calorimetry

ITC was attempted to assess Tyr binding to ^{WT}*Tma*DAH7PS and ^{V65P}*Tma*DAH7PS. ITC experiments were performed using a low volume Nano ITC (TA Instruments). ^{WT}*Tma*DAH7PS and ^{V65P}*Tma*DAH7PS samples were extensively dialysed into buffer (50 mM BTP, pH 7.3), which was also used to prepare Tyr solutions. Protein and Tyr solutions were degassed under vacuum for 10 minutes prior to use. The sample cell was loaded with 300 µL of protein sample (27-284 µM) and the syringe filled with 50 µL of Tyr solution (0.3-3 mM). 2 µL Tyr injections were performed at 240 second intervals. The heat of dilution was determined by injecting Tyr into buffer and subtracted from the experimental data. Experiments were performed at either 25 °C or 60 °C.

References

1. Davis, B. D., Intermediates in amino acid biosynthesis. *Advances in Enzymology and Related Subjects of Biochemistry* **1955,** *16*, 247-312.

2. Roberts, F.; Roberts, C. W.; Johnson, J. J.; Kyle, D. E.; Krell, T.; Coggins, J. R.; Coombs, G. H.; Milhous, W. K.; Tzipori, S.; Ferguson, D. J. P.; Chakrabarti, D.; McLeod, R., Evidence for the shikimate pathway in apicomplexan parasites. *Nature* **1998,** *393* (6687), 801-805.

3. Herrmann, K. M.; Weaver, L. M., The shikimate pathway. *Annu. Rev. Plant Physiol. Plant Molec. Biol.* **1999,** *50*, 473-503.

4. Davis, B. D., Aromatic biosynthesis .1. The role of shikimic acid. *J. Biol. Chem.* **1951,** *191* (1), 315-325.

5. Sprinson, D. B., The biosynthesis of aromatic compounds from D-glucose. *arbohydrate Chemistry* **1960,** *15*, 235-270.

6. Bentley, R., The shikimate pathway - a metabolic tree with many branches. *Critical Reviews in Biochemistry and Molecular Biology* **1990,** *25* (5), 307-384.

7. Tzin, V.; Galili, G., New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Molecular Plant* **2010,** *3* (6), 956-972.

8. Gibson, F.; Jackman, L. M., Structure of chorismic acid, a new intermediate in aromatic biosynthesis. *Nature* **1963,** *198* (487), 388.

9. Light, S. H.; Anderson, W. F., The diversity of allosteric controls at the gateway to aromatic amino acid biosynthesis. *Protein Sci.* **2013,** *22* (4), 395-404.

10. Tzin, V.; Galili, G., The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. *The Arabidopsis Book / American Society of Plant Biologists* **2010,** *8*, e0132.

11. Blackmore, N. J.; Reichau, S.; Jiao, W. T.; Hutton, R. D.; Baker, E. N.; Jameson, G. B.; Parker, E. J., Three sites and you are out: ternary synergistic allostery controls aromatic amino acid aiosynthesis in *Mycobacterium tuberculosis*. *J. Mol. Biol.* **2013,** *425* (9), 1582-1592.

12. Schonbrunn, E.; Eschenburg, S.; Shuttleworth, W. A.; Schloss, J. V.; Amrhein, N.; Evans, J. N.; Kabsch, W., Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. *Proc Natl Acad Sci U S A* **2001,** *98* (4), 1376-80.

13. Steinrucken, H. C.; Amrhein, N., The herbicide glyphosate is a potent nolpyruvyl-shikimic-acid 3-phophate synthase. *Biochem. Biophys. Res. Commun.* **1980,** *94* (4), 1207-1212.

14. Herrmann, K. M., The shikimate pathway - early steps in the biosynthesis of aromatic-compounds. *Plant Cell* **1995,** *7* (7), 907-919.

15. Maeda, H.; Dudareva, N., The shikimate pathway and aromatic amino acid biosynthesis in plants. In *Annual Review of Plant Biology, Vol 63*, Merchant, S. S., Ed. 2012; Vol. 63, pp 73-105.

16. Cross, P. J.; Allison, T. M.; Dobson, R. C. J.; Jameson, G. B.; Parker, E. J., Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110* (6), 2111-2116.

17. Jensen, R. A.; Nester, E. W., Regulatory significance of intermediary metabolites - control of aromatic amino acid biosynthesis by feedback inhibition in *Bacillus subtilis*. *J. Mol. Biol.* **1965,** *12* (2), 468.

18. Nester, E. W.; Jensen, R. A., Control of aromatic acid biosynthesis in Bacillus subtilis - sequential feedback inhibition. *Journal of Bacteriology* **1966,** *91* (4), 1594.

19. Pratelli, R.; Pilot, G., Regulation of amino acid metabolic enzymes and transporters in plants. *Journal of Experimental Botany* **2014,** *65* (19), 5535-5556.

20. Dev, A.; Tapas, S.; Pratap, S.; Kumar, P., Structure and function of enzymes of shikimate pathway. *Current Bioinformatics* **2012,** *7* (4), 374-391.

21. Staub, M.; Denes, G., A kinetic study of the mechanism of action of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase in *Escherichia coli* K 12. *Biochimica et biophysica acta* **1967,** *132* (2), 528-30.

22. Deleo, A. B.; Dayan, J.; Sprinson, D. B., Purification and kinetics of tyrosine-sensitive 3-Deoxy-D-arabino-heptulosonic acid 7-phosphate synthetase from salmonella. *J. Biol. Chem.* **1973,** *248* (7), 2344-2353.

23. Tran, D.; Pietersma, A. L.; Schofield, L. R.; Rost, M.; Jameson, G. B.; Parker, E. J., Investigating the role of the hydroxyl groups of substrate erythrose 4-phosphate in the reaction catalysed by the first enzyme of the shikimate pathway. *Bioorg. Med. Chem. Lett.* **2011,** *21* (22), 6838-6841.

24. Konig, V.; Pfeil, A.; Braus, G. H.; Schneider, T. R., Substrate and metal complexes of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Saccharomyces cerevisiae* provide new insights into the catalytic mechanism. *J. Mol. Biol.* **2004,** *337* (3), 675-690.

25. Sundaram, A. K.; Woodard, R. W., Probing the stereochemistry of *E-coli* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (phenylalanine-sensitive)-catalyzed synthesis of KDO 8-P analogues. *Journal of Organic Chemistry* **2000,** *65* (19), 5891-5897.

26. Lambert, J. M.; Boocock, M. R.; Coggins, J. R., The 3-dehydroquinate synthase activity of the pentafunctional arom enzyme complex of *Neurospora crassa* is Zn2+-dependent. *Biochem. J.* **1985,** *226* (3), 817-829.

27. Williamson, R. M.; Pietersma, A. L.; Jameson, G. B.; Parker, E. J., Stereospecific deuteration of 2-deoxyerythrose 4-phosphate using 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. *Bioorg. Med. Chem. Lett.* **2005,** *15* (9), 2339-2342.

28. Pilch, P. F.; Somerville, R. L., Fluorine-containing analogs of intermediates in the shikimate pathway. *Biochemistry* **1976,** *15* (24), 5315-5320.

29. Deleo, A. B.; Sprinson, D. B., Mechanism of 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthetase. *Biochem. Biophys. Res. Commun.* **1968,** *32* (5), 873.

30. Dotson, G. D.; Dua, R. K.; Clemens, J. C.; Wooten, E. W.; Woodard, R. W., Overproduction and one-step purification of *Escherichia coli* 3-deoxy-D-manno-octulosonic acid 8-phosphate synthase and oxygen-transfer studies during catalysis using isotopic-shifted heteronuclear NMR. *J. Biol. Chem.* **1995,** *270* (23), 13698-13705.

31. Howe, D. L.; Sundaram, A. K.; Wu, J.; Gatti, D. L.; Woodard, R. W., Mechanistic insight into 3-deoxy-D-manno-octulosonate-8-phosphate synthase and 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase utilizing phosphorylated monosaccharide analogues. *Biochemistry* **2003,** *42* (17), 4843-4854.

32. Jensen, R. A.; Xie, G.; Calhoun, D. H.; Bonner, C. A., The correct phylogenetic relationship of KdsA (3-deoxy-D-manno-octulosonate 8-phosphate synthase) with one of two independently evolved classes of AroA (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase). *Journal of Molecular Evolution* **2002,** *54* (3), 416-423.

33. Wu, J.; Howe, D. L.; Woodard, R. W., *Thermotoga maritima* 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase: the ancestral eubacterial DAHP synthase? *J. Biol. Chem.* **2003,** *278* (30), 27525-31.

34. Nazmi, A. R.; Schofield, L. R.; Dobson, R. C. J.; Jameson, G. B.; Parker, E. J., Destabilization of the homotetrameric assembly of 3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase from the hyperthermophile *Pyrococcus furiosus* enhances enzymatic activity. *J. Mol. Biol.* **2014,** *426* (3), 656-673.

35. Subramaniam, P. S.; Xie, G.; Xia, T. H.; Jensen, R. A., Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate synthase from *Neisseria gonorrhoeae* in the context of its membership in a protein family containing a subset of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases. *Journal of Bacteriology* **1998,** *180* (1), 119-127.

36. Gosset, G.; Bonner, C. A.; Jensen, R. A., Microbial origin of plant-type 2-keto-3-deoxy-D-arabino-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from Xanthomonas campestris. *J Bacteriol* **2001,** *183* (13), 4061-70.

37. Shumilin, I. A.; Bauerle, R.; Wu, J.; Woodard, R. W.; Kretsinger, R. H., Crystal structure of the reaction complex of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Thermotoga maritima* refines the catalytic mechanism and indicates a new mechanism of allosteric regulation. *J. Mol. Biol.* **2004,** *341* (2), 455-466.

38. Cross, P. J.; Pietersma, A. L.; Allison, T. M.; Wilson-Coutts, S. M.; Cochrane, F. C.; Parker, E. J., *Neisseria meningitidis* expresses a single 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase that is inhibited primarily by phenylalanine. *Protein Sci.* **2013,** *22* (8), 1087-1099.

39. Hartmann, M.; Schneider, T. R.; Pfeil, A.; Heinrich, G.; Lipscomb, W. N.; Braus, G. H., Evolution of feedback-inhibited beta/alpha barrel isoenzymes by gene duplication and a single mutation. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100* (3), 862-867.

40. Nagano, N.; Orengo, C. A.; Thornton, J. M., One fold with many functions: The evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.* **2002,** *321* (5), 741-765.

41. Webby, C. J.; Baker, H. M.; Lott, J. S.; Baker, E. N.; Parker, E. J., The structure of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *J. Mol. Biol.* **2005,** *354* (4), 927-939.

42. Stetter, K. O., Hyperthermophilic procaryotes. *Fems Microbiology Reviews* **1996,** *18* (2-3), 149-158.

43. Schofield, L. R.; Anderson, B. F.; Patchett, M. L.; Norris, G. E.; Jameson, G. B.; Parker, E. J., Substrate ambiguity and crystal structure of *Pyrococcus furiosus* 3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase: An ancestral 3-Deoxyald-2-ulosonate-phosphate synthase? *Biochemistry* **2005,** *44* (36), 11950-11962.

44. Zhou, L.; Wu, J.; Janakiraman, V.; Shumilin, I. A.; Bauerle, R.; Kretsinger, R. H.; Woodard, R. W., Structure and characterization of the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from *Aeropyrum pernix*. *Bioorganic Chem.* **2012,** *40*, 79-86.

45. Schofield, L. R.; Patchett, M. L.; Parker, E. J., Expression, purification, and characterization of 3-deody-D-arabino-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*. *Protein Expression and Purification* **2004,** *34* (1), 17-27.

46. Cross, P. J.; Parker, E. J., Allosteric inhibitor specificity of *Thermotoga maritima* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. *FEBS Lett.* **2013,** *587* (18), 3063-3068.

47. Liberles, J. S.; Thorolfson, M.; Martinez, A., Allosteric mechanisms in ACT domain containing enzymes involved in amino acid metabolism. *Amino Acids* **2005,** *28* (1), 1-12.

48. Cross, P. J.; Dobson, R. C. J.; Patchett, M. L.; Parker, E. J., Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *J. Biol. Chem.* **2011,** *286* (12), 10216-10224.

49. Light, S. H.; Halavaty, A. S.; Minasov, G.; Shuvalova, L.; Anderson, W. F., Structural analysis of a 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase with an N-terminal chorismate mutase-like regulatory domain. *Protein Sci.* **2012,** *21* (6), 887-895.

50. Wu, J.; Sheflyan, G. Y.; Woodard, R. W., *Bacillus subtilis* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase revisited: resolution of two long-standing enigmas. *Biochem. J.* **2005,** *390*, 583-590.

51. Simpson, R. J.; Davidson, B. E., Studies on 3-deoxy-D-arabinoheptulosonate-7-phosphate synthetase(phe)from *Escherichia coli* K12. 2. Kinetic properties. *European journal of biochemistry / FEBS* **1976,** *70* (2), 501-7.

52. Schoner, R.; Herrmann, K. M., 3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase. Purification, properties, and kinetics of the tyrosine-sensitive isoenzyme from *Escherichia coli*. *The Journal of biological chemistry* **1976,** *251* (18), 5440-7.

53. Shumilin, I. A.; Kretsinger, R. H.; Bauerle, R. H., Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Structure with Folding & Design* **1999,** *7* (7), 865-875.

54. Helmstaedt, K.; Strittmatter, A.; Lipscomb, W. N.; Braus, G. H., Evolution of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase-encoding genes in the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102* (28), 9784-9789.

55. Heyes, L. C.; Reichau, S.; Cross, P. J.; Jameson, G. B.; Parker, E. J., Structural analysis of substrate-mimicking inhibitors in complex with *Neisseria meningitidis* 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase - The importance of accommodating the active site water. *Bioorganic Chem.* **2014,** *57*, 242-250.

56. Jiao, W. T.; Hutton, R. D.; Cross, P. J.; Jameson, G. B.; Parker, E. J., Dynamic Cross-Talk among Remote Binding Sites: The molecular basis for unusual synergistic allostery. *J. Mol. Biol.* **2012,** *415* (4), 716-726.

57. Webby, C. J.; Jiao, W. T.; Hutton, R. D.; Blackmore, N. J.; Baker, H. M.; Baker, E. N.; Jameson, G. B.; Parker, E. J., Synergistic allostery, a sophisticated regulatory network for the control of aromatic amino acid biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2010,** *285* (40), 30567-30576.

58. Jeong, H.; Tombor, B.; Albert, R.; Oltvai, Z. N.; Barabasi, A. L., The large-scale organization of metabolic networks. *Nature* **2000,** *407* (6804), 651-654.

59. Sakami, W.; Harrington, H., Amino acid metabolism. *Annual review of biochemistry* **1963,** *32*, 355-98.

60. Braus, G. H., Aromatic amino acid biosynthesis in the yeast *Saccharomyces cerevisiae*: a model system for the regulation of a eukaryotic biosynthetic pathway. *Microbiological Reviews* **1991,** *55* (3), 349-370.

61. Kunzler, M.; Paravicini, G.; Egli, C. M.; Irniger, S.; Braus, G. H., Cloning, primary structure and regulation of the Aro4 gene, encoding the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Saccharomyces-cerevisiae*. *Gene* **1992,** *113* (1), 67-74.

62. Warden, S. M.; Richardson, C.; O'donnell, J.; Stapleton, D.; Kemp, B. E.; Witters, L. A., Post-translational modifications of the β-1 subunit of AMP-activated protein kinase affect enzyme activity and cellular localization. *Biochem. J.* **2001,** *354* (2), 275-283.

63. Nguyen, P. T. T.; Conley, A. J.; Sneyd, J.; Lee, R. S. F.; Soboleva, T. K.; Shorten, P. R., The role of enzyme compartmentalization on the regulation of steroid synthesis. *Journal of Theoretical Biology* **2013,** *332*, 52-64.

64. Shumilin, I. A.; Zhao, C.; Bauerle, R.; Kretsinger, R. H., Allosteric inhibition of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *J. Mol. Biol.* **2002,** *320* (5), 1147-1156.

65. Tsai, C. J.; del Sol, A.; Nussinov, R., Allostery: absence of a change in shape does not imply that allostery is not at play. *J Mol Biol* **2008,** *378* (1), 1-11.

66. Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J., The ensemble nature of allostery. *Nature* **2014,** *508* (7496), 331-339.

67. Goodey, N. M.; Benkovic, S. J., Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **2008,** *4* (8), 474-482.

68. Gunasekaran, K.; Ma, B.; Nussinov, R., Is allostery an intrinsic property of all dynamic proteins? *Proteins* **2004,** *57* (3), 433-43.

69. Monod, J.; Wyman, J.; Changeux, J. P., On nature of allosteric transitions - a plausible model. *J. Mol. Biol.* **1965,** *12* (1), 88.

70. Koshland, D. E.; Hamadani, K., Proteomics and models for enzyme cooperativity. *J. Biol. Chem.* **2002,** *277* (49), 46841-46844.

71. Kern, D.; Zuiderweg, E. R. P., The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol.* **2003,** *13* (6), 748-757.

72. Suel, G. M.; Lockless, S. W.; Wall, M. A.; Ranganathan, R., Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature Structural Biology* **2003,** *10* (1), 59-69.

73. Wu, J.; Woodard, R. W., New insights into the evolutionary links relating to the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase subfamilies. *J. Biol. Chem.* **2006,** *281* (7), 4042-4048.

74. Shumilin, I. A.; Bauerle, R.; Kretsinger, R. H., The high-resolution structure of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase reveals a twist in the plane of bound phosphoenolpyruvate. *Biochemistry* **2003,** *42* (13), 3766-3776.

75. Webby, C. J.; Patchett, M. L.; Parker, E. J., Characterization of a recombinant type II 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Helicobacter pylori*. *Biochem. J.* **2005,** *390*, 223-230.

76. Ma, B.; Tsai, C.-J.; Haliloğlu, T.; Nussinov, R., Dynamic allostery: linkers are not merely flexible. *Structure* **2011,** *19* (7), 907-917.

77. Woese, C. R.; Kandler, O.; Wheelis, M. L., Towards a natural system of organisms - proposal for the domains archaea, bacteria, and eucarya. *Proc. Natl. Acad. Sci. U. S. A.* **1990,** *87* (12), 4576-4579.

78. Huber, R.; Langworthy, T. A.; Konig, H.; Thomm, M.; Woese, C. R.; Sleytr, U. B.; Stetter, K. O., *Thermotoga-maritima* SP-NOV represents a new genus of unique extremely thermophilic eubacteria growing up to 90-degrees-c. *Archives of Microbiology* **1986,** *144* (4), 324-333.

79. Blamey, J. M.; Adams, M. W. W., Characterization of an ancestral type of pyruvate ferredoxin oxidoreductase from the hyperthermophilic bacterium, *Thermotoga-maritima*. *Biochemistry* **1994,** *33* (4), 1000-1007.

80. Latif, H.; Lerman, J. A.; Portnoy, V. A.; Tarasova, Y.; Nagarajan, H.; Schrimpe-Rutledge, A. C.; Smith, R. D.; Adkins, J. N.; Lee, D.-H.; Qiu, Y.; Zengler, K., The genome organization of *Thermotoga maritima* reflects its lifestyle. *PLoS Genetics* **2013,** *9* (4), e1003485.

81. Ma, B.; Tsai, C.-J.; Haliloğlu, T.; Nussinov, R., Dynamic allostery: linkers are not merely flexible. *Structure 19* (7), 907-917.

82. Cross, P. J., Unravelling the evolution of allosteric regulation in 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate Synthase. Doctoral thesis, University of Canterbury, New Zealand, **2012**.

83. Tabor, S.; Richardson, C. C., A Bacteriophage-T7 RNA-polymerase promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A.* **1985,** *82* (4), 1074-1078.

84. Niesen, F. H.; Berglund, H.; Vedadi, M., The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols* **2007,** *2* (9), 2212-21.

85. Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I., PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* **2003,** *36*, 1277-1282.

86. Svergun, D.; Barberato, C.; Koch, M. H. J., CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography* **1995,** *28*, 768-773.

87. Tzeng, S. R.; Kalodimos, C. G., Protein dynamics and allostery: an NMR view. *Curr. Opin. Struct. Biol.* **2011,** *21* (1), 62-67.

88. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005,** *26* (16), 1781-1802.

89. Boehr, D. D.; Nussinov, R.; Wright, P. E., The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **2009,** *5* (11), 789-796.

90. Wrabl, J. O.; Gu, J. N.; Liu, T.; Schrank, T. P.; Whitten, S. T.; Hilser, V. J., The role of protein conformational fluctuations in allostery, function, and evolution. *Biophysical Chemistry* **2011,** *159* (1), 129-141.

91. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-972.

92. Bahar, I.; Chennubhotla, C.; Tobi, D., Intrinsic dynamics of enzymes in the unbound state and, relation to allosteric regulation. *Curr. Opin. Struct. Biol.* **2007,** *17* (6), 633-640.

93. Tsai, C. J.; Del Sol, A.; Nussinov, R., Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Molecular Biosystems* **2009,** *5* (3), 207-216.

94. Lang, E. J. M.; Cross, P. J.; Mittelstadt, G.; Jameson, G. B.; Parker, E. J., Allosteric ACTion: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr. Opin. Struct. Biol.* **2014,** *29*, 102-111.

95. Tuckerman, M. E.; Martyna, G. J., Understanding modern molecular dynamics: Techniques and applications. *Journal of Physical Chemistry B* **2000,** *104* (2), 159-178.

96. Ponder, J. W.; Case, D. A., Force fields for protein simulations. *Protein Simulations* **2003,** *66*, 27.

97. Makov, G.; Payne, M. C., Periodic boundary conditions in ab-*initio* calculations. *Physical Review B* **1995,** *51* (7), 4014-4022.

98. Hansson, T.; Oostenbrink, C.; van Gunsteren, W. F., Molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2002,** *12* (2), 190-196.

99. Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A., Molecular dynamics simulations on solvated biomolecular systems: The particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* **1995,** *117* (14), 4193-4194.

100. Lang, E. J. M., Unpublished doctoral thesis. University of Canterbury, Christchurch, New Zealand, **2016**.

101. Ferrara, P.; Apostolakis, J.; Caflisch, A., Targeted molecular dynamics simulations of protein unfolding. *Journal of Physical Chemistry B* **2000,** *104* (18), 4511-4518.

102. Schlitter, J.; Engels, M.; Kruger, P., Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *Journal of molecular graphics* **1994,** *12* (2), 84-9.

103. Ovchinnikov, V.; Karplus, M., Analysis and elimination of a bias in targeted molecular dynamics simulations of conformational transitions: Application to calmodulin. *The Journal of Physical Chemistry. B* **2012,** *116* (29), 8584-8603.

104. Liu, J.; Nussinov, R., Molecular dynamics reveal the essential role of linker motions in the function of cullin-RING E3 ligases. *J. Mol. Biol.* **2010,** *396* (5), 1508-1523.

105. Akimoto, M.; Moleschi, K.; Boulton, S.; VanSchouwen, B.; Selvaratnam, R.; Taylor, S. S.; Melacini, G., Allosteric linkers in cAMP signalling. *Biochemical Society Transactions* **2014,** *42*, 139-144.

106. Swain, J. F.; Dinler, G.; Sivendran, R.; Montgomery, D. L.; Stotz, M.; Gierasch, L. M., Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Molecular Cell* **2007,** *26* (1), 27-39.

107. Akimoto, M.; Selvaratnam, R.; McNicholl, E. T.; Verma, G.; Taylor, S. S.; Melacini, G., Signaling through dynamic linkers as revealed by PKA. *Proc Natl Acad Sci U S A* **2013,** *110* (35), 14231-6.

108. Pierce, M. M.; Raman, C. S.; Nall, B. T., Isothermal titration calorimetry of protein-protein interactions. *Methods-a Companion to Methods in Enzymology* **1999,** *19* (2), 213-221.

109. Schrödinger; LLC, The PyMOL Molecular Graphics System, Version 1.8 **2015**.

110. McWilliam H, L. W., Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R,; Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res. 41(Web Server issue)*, 597-600.

111. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **2004,** *60*, 2126-2132.

112. Song, Y.; Mao, J.; Gunner, M. R., MCCE2: improving protein pKa calculations with extensive side chain rotamer sampling. *J Comput Chem* **2009,** *30* (14), 2231-47.

113. Olsson, M. H.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of chemical theory and computation* **2011,** *7* (2), 525-37.

114. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics & Modelling* **1996,** *14* (1), 33-38.

115. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics (Oxford, England)* **2013,** *29* (7), 845-54.